

A level playing field?

Drug testing in sport aims to promote fair play, but the science behind the tests needs to be more open.

The opening ceremony of the 2008 Olympic Games in Beijing has yet to take place, but already several athletes face charges of taking substances banned by the International Olympic Committee. These are just the latest in a long line of cases in which competitors have been accused of using performance-enhancing substances. Far from quelling such practices, the advent of drug testing in sport in the late 1960s stimulated an arms race between regulators and the cheats.

Today, some athletes and their coaches continue to risk their reputation, and sometimes the athletes' long-term health, for the chance to dope undetected. In the process, they push the human body to its limits and go beyond what is known about the drugs being taken. The latest drugs are designed with testing in mind, so that they either clear from the body quickly or do not produce the tell-tale metabolite spikes in blood and urine samples. As a result, the testing labs must also push to stay one step ahead of the cheats.

On page 692, biostatistician Donald Berry of the University of Texas in Houston outlines what he sees as problems with the way doping tests are conducted. He argues that anti-doping authorities have not adequately defined and publicized how they arrived at the criteria used to determine whether or not a test result is positive. The ability of an anti-doping test to detect a banned substance in an athlete is calibrated

in part by testing a small number of volunteers taking the substance in question. But Berry says that individual labs need to verify these detection limits in larger groups that include known dopers and non-dopers under blinded conditions that mimic what happens during competition.

Nature believes that accepting 'legal limits' of specific metabolites without such rigorous verification goes against the foundational standards of modern science, and results in an arbitrary test for which the rate of false positives and false negatives can never be known. By leaving these rates unknown, and by not publishing and opening to broader scientific scrutiny the methods by which testing labs engage in study, it is *Nature's* view that the anti-doping authorities have fostered a sporting culture of suspicion, secrecy and fear.

Detecting cheats is meant to promote fairness, but drug testing should not be exempt from the scientific principles and standards that apply to other biomedical sciences, such as disease diagnostics. The alternative could see the innocent being punished while the guilty escape on the grounds of reasonable doubt. ■

"Drug testing should not be exempt from the scientific principles and standards that apply to other biomedical sciences."

Clean hands, please

The Italian government needs to maintain a careful distance from industry.

Fifteen years ago, at the height of Italy's 'Clean Hands' anti-corruption campaign, police broke into the house of Duilio Poggiolini, head of the national committee for drug registration, and discovered gold bullion under his floorboards. For many Italians, the image of that gleaming bullion still resonates — an enduring symbol of a time when government officials, up to and including the health minister, routinely took bribes from the pharmaceutical industry to approve drugs and fix their prices.

Steps were taken to avoid such a situation arising again. So it is worrying that Nello Martini, a pharmacist with no political associations, has been removed by Prime Minister Silvio Berlusconi's new government as head of AIFA, the autonomous agency created in 2004 to register drugs and supervise their use. Martini successfully carried out a mandate to limit spiralling drug expenditure to 13% of the total health budget. But in the process he incurred the wrath of industry. Only a few weeks ago, government prosecutors in Turin charged Martini with *disastro colposo*, or 'causing unintentional disaster', for bureaucratic delays in updating the packaging information on the side effects of a few drugs — although none

required more than minor rewording of existing text.

Martini was replaced in the middle of July by microbiologist Guido Rasi, a member of AIFA's administrative board, who has been described in the Italian press as being close to the far-right party Alleanza Nazionale, which forms part of Berlusconi's coalition government. Even more worryingly, the government, which took office in May, says it plans to reduce AIFA's power by separating the pricing of drugs from technical considerations of their efficacy, bringing pricing back into the health and welfare ministry.

At a time when all countries are struggling to find a way to pay for hugely expensive new-generation drugs within limited budgets, this makes little sense. The autonomous agency needs to be able to integrate all technical and economic information if Italy is to operate a cost-effective health system. Moreover, the health and welfare ministry's connections with industry are uncomfortably close. For example, the wife of the minister Maurizio Sacconi is the director-general of Farminindustria, the association that promotes the interests of the pharmaceutical industry.

In fact, Berlusconi's government has shown unsettling tendencies to allow industrial interests to gain influence over state agencies. A few weeks after Martini's dismissal, the Italian space agency was put into the hands of a commissioner who heads the space division of the aerospace giant Finmeccanica. The government should think twice about whether it really wants to open the door that was deliberately closed after the Poggiolini affair. ■

RESEARCH HIGHLIGHTS

Turf wars

Global Change Biol. doi:10.1111/j.1365-2486.2008.01617.x (2008)

Lawns are already under environmental scrutiny owing to the amount of water they consume. But the emission of nitrous oxide (N_2O), an important greenhouse gas, might also be a problem.

Diane Pataki at the University of California, Irvine, and her colleagues used fertilized and unfertilized experimental turf plots heated by about $3.5^\circ C$ and compared these with unheated controls. They show that warmer and wetter conditions increase N_2O emissions from fertilized lawns. As temperatures rise, they argue, turf fertilization will become a significant source of N_2O in urban areas.

Warmer conditions also promoted the growth of weeds with C_4 metabolism — which is more efficient than C_3 at higher temperatures — in the C_3 turf, adding to ongoing concerns about how weeds might spread under climate change.



BURKE/TRIOLO PRODUCTIONS/PHOTOLIBRARY.COM

ASTROPHYSICS

First light

Science **321**, 669–671 (2008)

Observations of distant quasars and galaxies show that there was light when the Universe was less than one billion years old. But how did the first stars coalesce and ignite their nuclear fire?

Naoki Yoshida of Nagoya University in Japan and his colleagues used supercomputers to simulate the formation of the gravitational seeds for the first stars, beginning with the hot gas and cold dark matter right after the Big Bang and following their gravitational collapse. The models spanned 13 orders of magnitude, tracking the complicated motions of volumes of gas smaller than the Sun amid a primeval cosmic medium hundreds of thousands of light years across.

The resulting protostars eventually evolved into massive stars, a hundred times the mass of the Sun.

ECOLOGY

Fungus hunters

Naturwissenschaften doi:10.1007/s00114-008-0421-9 (2008)

Ants have developed many survival strategies, from seed harvesting to fungus gardening. But the *Euprenolepis procera* ant (pictured right) of southeast Asia has an entirely new strategy: gathering wild mushrooms.

Volker Witte of Ludwig-Maximilians University in Munich, Germany, and Ulrich Maschwitz of Johann-Wolfgang-Goethe University in Frankfurt tracked the ants in the Malaysian rainforest and discovered that

they are both nomadic and can survive on mushrooms alone. Although some other ants are known to be nomads, the diet of wild mushrooms is a surprise, as they are often loaded with toxins and low in nutrients.

The authors argue that once the ability to consume fungus developed, intense competition with other species drove them to devote themselves to this food specially.

ANIMAL BEHAVIOUR

Is he into her?

Curr. Biol. doi:10.1016/j.cub.2008.06.067 (2008)

When it comes to choosing a mate, male Atlantic molly fish (*Poecilia mexicana*) often have to make their selection when other males are watching. Martin Plath at the University of Potsdam in Germany and his colleagues suggest that, in this situation, males try to deceive others about their true intentions regarding mate choice.

When given their choice of a small or large female, males without an audience usually made advances to the bigger one. When another male was present, however, the

molly initially paid attention to the smaller female but then stopped expressing a mating preference.

Plath hypothesizes that the behaviour is a counteradaptation to male mate-choice copying, in which males try to mate with a female that others also pursued.

PALAEOCLIMATE

Quick start to a cold spell

Nature Geosci. doi:10.1038/ngeo263 (2008)

During the winter of 12,679 years ago, western Europe was apparently slammed with a major wind shift that heralded the start of the coldest period in recent history.

Earlier work suggested that the Younger Dryas cold spell began rapidly around 12,700 years ago. Now, a team led by Achim Brauer of the GFZ German Research Centre for Geosciences in Potsdam, Germany, has narrowed this down by studying sediments laid down annually in a German lake.

The layers, called varves, record a dramatic shift in storminess over the course of a single year. The authors argue that this reflects changes in atmospheric circulation patterns over the Atlantic, and helps explain how a shutdown of ocean circulation could have led to abrupt cooling in western Europe.

MOLECULAR BIOLOGY

Going farther

Cell doi:10.1016/j.cell.2008.06.051 (2008)

Researchers have identified what they say are the first targeted drugs that boost physical endurance.

The team, led by Ronald Evans at the Salk Institute in La Jolla, California, tested



RESEARCH HIGHLIGHTS

Turf wars

Global Change Biol. doi:10.1111/j.1365-2486.2008.01617.x (2008)

Lawns are already under environmental scrutiny owing to the amount of water they consume. But the emission of nitrous oxide (N_2O), an important greenhouse gas, might also be a problem.

Diane Pataki at the University of California, Irvine, and her colleagues used fertilized and unfertilized experimental turf plots heated by about 3.5°C and compared these with unheated controls. They show that warmer and wetter conditions increase N_2O emissions from fertilized lawns. As temperatures rise, they argue, turf fertilization will become a significant source of N_2O in urban areas.

Warmer conditions also promoted the growth of weeds with C_4 metabolism — which is more efficient than C_3 at higher temperatures — in the C_3 turf, adding to ongoing concerns about how weeds might spread under climate change.



BURKE/TRIOLO PRODUCTIONS/PHOTOLIBRARY.COM

ASTROPHYSICS

First light

Science **321**, 669–671 (2008)

Observations of distant quasars and galaxies show that there was light when the Universe was less than one billion years old. But how did the first stars coalesce and ignite their nuclear fire?

Naoki Yoshida of Nagoya University in Japan and his colleagues used supercomputers to simulate the formation of the gravitational seeds for the first stars, beginning with the hot gas and cold dark matter right after the Big Bang and following their gravitational collapse. The models spanned 13 orders of magnitude, tracking the complicated motions of volumes of gas smaller than the Sun amid a primeval cosmic medium hundreds of thousands of light years across.

The resulting protostars eventually evolved into massive stars, a hundred times the mass of the Sun.

ECOLOGY

Fungus hunters

Naturwissenschaften doi:10.1007/s00114-008-0421-9 (2008)

Ants have developed many survival strategies, from seed harvesting to fungus gardening. But the *Euprenolepis procera* ant (pictured right) of southeast Asia has an entirely new strategy: gathering wild mushrooms.

Volker Witte of Ludwig-Maximilians University in Munich, Germany, and Ulrich Maschwitz of Johann-Wolfgang-Goethe University in Frankfurt tracked the ants in the Malaysian rainforest and discovered that

they are both nomadic and can survive on mushrooms alone. Although some other ants are known to be nomads, the diet of wild mushrooms is a surprise, as they are often loaded with toxins and low in nutrients.

The authors argue that once the ability to consume fungus developed, intense competition with other species drove them to devote themselves to this food specially.

ANIMAL BEHAVIOUR

Is he into her?

Curr. Biol. doi:10.1016/j.cub.2008.06.067 (2008)

When it comes to choosing a mate, male Atlantic molly fish (*Poecilia mexicana*) often have to make their selection when other males are watching. Martin Plath at the University of Potsdam in Germany and his colleagues suggest that, in this situation, males try to deceive others about their true intentions regarding mate choice.

When given their choice of a small or large female, males without an audience usually made advances to the bigger one. When another male was present, however, the

molly initially paid attention to the smaller female but then stopped expressing a mating preference.

Plath hypothesizes that the behaviour is a counteradaptation to male mate-choice copying, in which males try to mate with a female that others also pursued.

PALAEOCLIMATE

Quick start to a cold spell

Nature Geosci. doi:10.1038/ngeo263 (2008)

During the winter of 12,679 years ago, western Europe was apparently slammed with a major wind shift that heralded the start of the coldest period in recent history.

Earlier work suggested that the Younger Dryas cold spell began rapidly around 12,700 years ago. Now, a team led by Achim Brauer of the GFZ German Research Centre for Geosciences in Potsdam, Germany, has narrowed this down by studying sediments laid down annually in a German lake.

The layers, called varves, record a dramatic shift in storminess over the course of a single year. The authors argue that this reflects changes in atmospheric circulation patterns over the Atlantic, and helps explain how a shutdown of ocean circulation could have led to abrupt cooling in western Europe.

MOLECULAR BIOLOGY

Going farther

Cell doi:10.1016/j.cell.2008.06.051 (2008)

Researchers have identified what they say are the first targeted drugs that boost physical endurance.

The team, led by Ronald Evans at the Salk Institute in La Jolla, California, tested



the effects of two drugs. The first, called GW1516, boosts activity of the gene that encodes PPAR δ , a protein involved in metabolism, and, when combined with exercise training, allowed mice to run up to 70% farther than untreated mice.

The researchers then dosed mice with the second drug, AICAR, which acts on an enzyme known as AMPK that is normally activated by exercise. Even though the mice did not undergo exercise training, AICAR boosted their endurance by 44% and remodelled their physiological and genetic characteristics in ways similar to exercise.

ANIMAL BEHAVIOUR

Love song

Behav. Ecol. Sociobiol. **62**, 1633–1641 (2008)

Most male birds sing to attract females, but some species have puzzled ornithologists by continuing to sing long after egg laying.

Valentin Amrhein and his colleagues at the University of Oslo compared blue tits (*Cyanistes caeruleus*) and great tits (*Parus major*). Both live in the same places and have similar reproductive behaviour. They differ in that the monogamous great tits often raise second broods; blue tits seldom do.

The team found that great tits carried on singing after egg laying, whereas blue tits reduced singing. They argue that great tits sing both to continue defending their territory and to encourage females to lay a second clutch.

CELL BIOLOGY

Starve and reproduce

Aging Cell doi: 10.1111/j.1474-9726.2008.00409.x (2008)

Mice that spend much of their adult lives on a restricted diet lengthen their fertile lifespan, a new study suggests.

Previous studies have shown that caloric restriction from birth can extend fertility, but can also adversely affect development. A team led by Jonathan Tilly at Massachusetts General Hospital in Boston cut animals' calorie intakes by 40% between 4 and 15.5 months of age — roughly equivalent to a human's mid-20s to mid-40s.

When returned to a normal diet, these mice continued to have young for several months beyond the time that they would normally stop reproducing. Moreover, many more of the offspring survived than did those born to control mice during the normal fertile period.

The work supports the idea that drugs mimicking caloric restriction might delay human menopause.

NEUROBIOLOGY

Baby blues

Neuron **59**, 207–213 (2008)

Levels of neurosteroids derived from the hormone progesterone fluctuate during the menstrual cycle and soar during pregnancy. But abnormal levels are associated with disorders such as premenstrual dysphoric syndrome and postpartum depression.

Neurosteroids act through the GABA_A receptor to modulate excitability of brain cells. Using genetically modified mice that lack the delta subunit of the GABA_A receptor, Istvan Mody and Jamie Maguire at the University of California, Los Angeles, found that females showed depressive and anxious behaviours after giving birth, and did not care properly for their pups.

Failure to regulate GABA_A receptors as

neurosteroid levels plummet after giving birth may similarly underlie human postpartum depression.

GEOLOGY

The end of flat Earth

Geology **36**, 635–638 (2008)

Dramatic mountains couldn't appear on Earth until its outer shell was strong enough to sustain them. New findings indicate this may have occurred during the Neoproterozoic era, between 2.8 billion and 2.5 billion years ago.



Numerical simulations run by Patrice Rey at the University of Sydney in Australia and Nicolas Coltice at the University of Lyon in France suggest that only in the Neoproterozoic did the Earth cool enough for its lithosphere to strengthen. That, in turn, allowed the crust to thicken and be uplifted to heights greater than 2,500 metres for the first time.

Newly rising mountains would have also meant more high elevations at which erosion could occur, which would have altered the geochemical links between planet and atmosphere.

P. BOURSEILLER/GETTY

JOURNAL CLUB

Benny Freeman
University of Texas at Austin

A chemical engineer notes that not all membrane pores are made equal; some are more equal than others.

Few cheap, man-made membranes have holes of uniform size. This makes them either inefficient or unreliable sieves of particles such as viruses. But membranes are also one of the least energy-intensive separation devices. As fuel costs rise, many of the billion or so people

without access to safe drinking water find it harder to sanitize what water they have. This is why I view the low-cost manufacture of isoporous membranes as a holy grail in the field.

Recently, some scientists in Germany unearthed a path to this chalice by tinkering with a technique known as 'phase inversion'. This is often used to make synthetic membranes: a polymer solution is immersed in a liquid, often water, which diffuses into the solution and causes a thin, porous membrane of hydrophobic polymer to form. The solid polymer

is a twisted, irregular matrix, full of odd-shaped pores.

Klaus-Viktor Peinemann and his co-workers started with a polymer in which the chain-like molecules have a hydrophobic and a hydrophilic end, and allowed the solvent solution to evaporate. As this happened, they think that the polymer assembled into connected cylinders, with the hydrophobic and hydrophilic parts of different molecules coming together. The researchers then plunged this nascent membrane into water, which moved through the hydrophilic cylinders, opening

them up and thus creating identical and aligned pores (K.-V. Peinemann *et al. Nature Mater.* **6**, 992–996; 2007).

The pores were all about 10 nanometres wide — roughly the right size to separate hepatitis B virus from water. Picking other polymers with hydrophobic and hydrophilic parts should allow the development of membranes with uniform-diameter pores of various sizes. That could be a boon for industry as well as public health.

Discuss this paper at <http://blogs.nature.com/nature/journalclub>

the effects of two drugs. The first, called GW1516, boosts activity of the gene that encodes PPAR δ , a protein involved in metabolism, and, when combined with exercise training, allowed mice to run up to 70% farther than untreated mice.

The researchers then dosed mice with the second drug, AICAR, which acts on an enzyme known as AMPK that is normally activated by exercise. Even though the mice did not undergo exercise training, AICAR boosted their endurance by 44% and remodelled their physiological and genetic characteristics in ways similar to exercise.

ANIMAL BEHAVIOUR

Love song

Behav. Ecol. Sociobiol. **62**, 1633–1641 (2008)

Most male birds sing to attract females, but some species have puzzled ornithologists by continuing to sing long after egg laying.

Valentin Amrhein and his colleagues at the University of Oslo compared blue tits (*Cyanistes caeruleus*) and great tits (*Parus major*). Both live in the same places and have similar reproductive behaviour. They differ in that the monogamous great tits often raise second broods; blue tits seldom do.

The team found that great tits carried on singing after egg laying, whereas blue tits reduced singing. They argue that great tits sing both to continue defending their territory and to encourage females to lay a second clutch.

CELL BIOLOGY

Starve and reproduce

Aging Cell doi: 10.1111/j.1474-9726.2008.00409.x (2008)

Mice that spend much of their adult lives on a restricted diet lengthen their fertile lifespan, a new study suggests.

Previous studies have shown that caloric restriction from birth can extend fertility, but can also adversely affect development. A team led by Jonathan Tilly at Massachusetts General Hospital in Boston cut animals' calorie intakes by 40% between 4 and 15.5 months of age — roughly equivalent to a human's mid-20s to mid-40s.

When returned to a normal diet, these mice continued to have young for several months beyond the time that they would normally stop reproducing. Moreover, many more of the offspring survived than did those born to control mice during the normal fertile period.

The work supports the idea that drugs mimicking caloric restriction might delay human menopause.

NEUROBIOLOGY

Baby blues

Neuron **59**, 207–213 (2008)

Levels of neurosteroids derived from the hormone progesterone fluctuate during the menstrual cycle and soar during pregnancy. But abnormal levels are associated with disorders such as premenstrual dysphoric syndrome and postpartum depression.

Neurosteroids act through the GABA_A receptor to modulate excitability of brain cells. Using genetically modified mice that lack the delta subunit of the GABA_A receptor, Istvan Mody and Jamie Maguire at the University of California, Los Angeles, found that females showed depressive and anxious behaviours after giving birth, and did not care properly for their pups.

Failure to regulate GABA_A receptors as

neurosteroid levels plummet after giving birth may similarly underlie human postpartum depression.

GEOLOGY

The end of flat Earth

Geology **36**, 635–638 (2008)

Dramatic mountains couldn't appear on Earth until its outer shell was strong enough to sustain them. New findings indicate this may have occurred during the Neoproterozoic era, between 2.8 billion and 2.5 billion years ago.



Numerical simulations run by Patrice Rey at the University of Sydney in Australia and Nicolas Coltice at the University of Lyon in France suggest that only in the Neoproterozoic did the Earth cool enough for its lithosphere to strengthen. That, in turn, allowed the crust to thicken and be uplifted to heights greater than 2,500 metres for the first time.

Newly rising mountains would have also meant more high elevations at which erosion could occur, which would have altered the geochemical links between planet and atmosphere.

P. BOURSEILLER/GETTY

JOURNAL CLUB

Benny Freeman
University of Texas at Austin

A chemical engineer notes that not all membrane pores are made equal; some are more equal than others.

Few cheap, man-made membranes have holes of uniform size. This makes them either inefficient or unreliable sieves of particles such as viruses. But membranes are also one of the least energy-intensive separation devices. As fuel costs rise, many of the billion or so people

without access to safe drinking water find it harder to sanitize what water they have. This is why I view the low-cost manufacture of isoporous membranes as a holy grail in the field.

Recently, some scientists in Germany unearthed a path to this chalice by tinkering with a technique known as 'phase inversion'. This is often used to make synthetic membranes: a polymer solution is immersed in a liquid, often water, which diffuses into the solution and causes a thin, porous membrane of hydrophobic polymer to form. The solid polymer

is a twisted, irregular matrix, full of odd-shaped pores.

Klaus-Viktor Peinemann and his co-workers started with a polymer in which the chain-like molecules have a hydrophobic and a hydrophilic end, and allowed the solvent solution to evaporate. As this happened, they think that the polymer assembled into connected cylinders, with the hydrophobic and hydrophilic parts of different molecules coming together. The researchers then plunged this nascent membrane into water, which moved through the hydrophilic cylinders, opening

them up and thus creating identical and aligned pores (K.-V. Peinemann *et al. Nature Mater.* **6**, 992–996; 2007).

The pores were all about 10 nanometres wide — roughly the right size to separate hepatitis B virus from water. Picking other polymers with hydrophobic and hydrophilic parts should allow the development of membranes with uniform-diameter pores of various sizes. That could be a boon for industry as well as public health.

Discuss this paper at <http://blogs.nature.com/nature/journalclub>

NEWS

Death renews biosecurity debate

The suicide of a biodefence researcher who was being investigated in connection with the 2001 anthrax attacks has raised questions about the US government's regulation of research on dangerous pathogens — even as Congress considers a bill to improve oversight.

The Federal Bureau of Investigation (FBI) had been working to link Bruce Ivins, a microbiologist at the US Army Medical Research Institute of Infectious Diseases (USAMRIID) in Frederick, Maryland, to five people who died after anthrax spores were sent through the mail soon after the terrorist attacks on 11 September 2001. The case has still not been solved, some say because of a long-running focus on a single, different suspect (see 'How the anthrax investigation unfolded').

Ivins died on 29 July of an overdose of Tylenol, a paracetamol-based painkiller. His lawyer and many of his colleagues maintain his innocence, and the FBI's evidence allegedly implicating Ivins had not been released as *Nature* went to press.

"If it turns out the evidence is compelling, it does raise a lot of questions about the insider threat and whether the current vetting procedure is valid," says Jonathan Tucker, a bioweapons expert with the Commission on the Prevention of Weapons of Mass Destruction Proliferation and Terrorism based in the Washington DC area.

Although the FBI had been investigating Ivins for months, he continued working at USAMRIID until 10 July this year. A social worker had reported him as being mentally unstable, and some co-workers have said the pressure of the investigation had got to him.

In 2002, Ivins was involved in a security breach when he failed to report contamination outside a biosecure laboratory to his supervisors. But officials apparently let the incident go, afraid that punishing Ivins would discourage others from reporting contamination.

Ivins was one of about 14,000 US researchers with clearance to work with pathogens or toxins from the government's list of 72 'select agents', such as the Ebola virus, avian influenza and ricin. These scientists must pass an FBI security assessment that includes fingerprinting, and checking for a criminal record or any history of drug abuse or hospitalization for mental illness. Congress is considering a bill that would, among other things, have the National Academy of Sciences evaluate the impact so far of the select-agent programme.

Working with select agents has required more

paperwork since the passage of the Patriot Act of 2001 and the Bioterrorism Act of 2002. "A lot of good scientists have said it's just too much trouble," Tucker says.

The penalties for running afoul of the regulations can be strict. "If you forget to file some of the appropriate paperwork, or some of the vials in your inventory get autoclaved and nobody writes it down, you can get in trouble," says James Roth, a professor of veterinary medicine at Iowa State University in Ames. In 2004, Thomas Butler, a microbiologist at Texas Tech University in Lubbock, received a two-year prison sentence in a case that started with missing plague samples.

Yet the increased interest in research on bioterror agents has been a boon for some. Civilian spending on biodefence rose from \$685 million in 2001 to a peak of \$8.2 billion in 2005, before levelling off at around \$5 billion in recent years.

"The research I'm doing on anthrax spores is work that I would have been interested in doing in other species, but I chose to do it in anthrax because there was plenty of funding available," says David Popham, a microbiologist at Virginia Polytechnic Institute and State University in Blacksburg. Because he works with a harmless variant of anthrax, he did not have to go through the select-agent clearance procedure.

In his research Ivins prepared spores of a virulent strain of *Bacillus anthracis* and infected



Bruce Ivins was to be indicted in connection with the anthrax attacks.

lab animals with them, to evaluate new vaccines and treatments. He co-authored, for example, a report published this week that tested whether blood plasma from humans treated with an anthrax vaccine could protect mice from the pathogen (J. F. Hewetson *et al.* *Vaccine* 26, 4262–4266; 2008). Ivins worked at the Army lab for more than 35 years, and received the Department of Defense's highest civilian honour in 2003 for his work

on an anthrax vaccine.

The fallout for the rest of the research community is not clear. "Are all of our biomedical scientists going to have to undergo evaluation deep into their psyches?" asks Alan Pearson, director of the Biological and Chemical Weapons Control Program at the Center for Arms Control and Non-Proliferation in Washington DC. "I think the opposition is going to be fierce."

Tara O'Toole, director of the Center for Biosecurity at the University of Pittsburgh Medical Center in Baltimore, Maryland, says that it is impossible to secure all pathogens because they occur in nature. "The notion that we can somehow prevent a bioattack by locking up pathogens in research laboratories is ridiculous," she says. Instead, she says, the solution is to be found in better medicines and vaccines. "We have to make it so hard to kill a lot of people that the terrorists aren't interested in trying."

Amber Dance

For more on this story, see <http://tinyurl.com/66scnf>.

HOW THE ANTHRAX INVESTIGATION UNFOLDED

September–October 2001

Letters containing anthrax spores are mailed from New Jersey to reporters in New York and Florida, as well as members of Congress in Washington DC. Five people die from anthrax exposure.

August 2002

Attorney-general John Ashcroft names bioweapons expert Steven Hatfill as a "person of interest" in the case.

August 2003

Hatfill sues the Department of Justice and, later, media outlets, for mentioning his name in connection with the investigation.

August 2004

The FBI searches the home of New York physician Kenneth Berry, who had worked on emergency responses to bioterrorism. No charges are brought.

June 2008

The Department of Justice agrees to pay Hatfill \$5.85 million to drop his civil suit.

July 2008

Bruce Ivins, a microbiologist who had worked at the US Army Medical Research Institute of Infectious Diseases, commits suicide in Frederick, Maryland. The FBI had been investigating him in relation to the 2001 attacks.

USAMRIID

Climate war games

More than 40 negotiators from Asia, Europe and the United States converged on Washington DC last week for what was billed as the first major war game involving global warming.

The Center for a New American Security, a Washington-based national-security think tank, gathered together climate scientists and experts in security, environmental policy and business for the role-playing exercise. Each was assigned to one of four teams, representing Europe, the United States, China and India, and charged with negotiating the best deal for their team.

Under the scenario, set in 2015, greenhouse-gas emissions are rising, and the latest climate models paint a grim picture of the future if business were to continue as usual. Extreme weather, including droughts, storms and floods, is on the rise. The United Nations is calling for international cooperation on emissions reductions, adaptation, disaster relief and shortages of crucial resources such as food and water.

Conference organizers worked with scientists at Oak Ridge National Laboratory in Tennessee who provided a climate simulation up to 2100 based on the worst-case scenario proposed by the Intergovernmental Panel on Climate Change. This scenario mixes rapid growth with continued reliance on fossil fuels — something the organizers say is reasonable, given that emissions are currently trending higher than projected.

After lengthy negotiations, the United States and the Europe Union agreed to a 30% reduction in emissions by 2025. They also agreed to finance a portion of the emissions reductions by the developing world. But China refused to accept any specific emissions targets. India committed itself to reductions, but only with a number of contingencies. In a sad mirror of reality, the participants departed after three days without the comprehensive agreement many had been hoping for.

“If it were easy it wouldn’t be realistic, because these are not easy issues,” says Reid Detchon, a US team member and executive director for energy and climate at the United Nations Foundation in Washington DC. Detchon credits the exercise with “exposing some of the tensions” in the debate — particularly on the China team, which included members from China.

“I was just struck by the enormity of the challenge before us,” says Todd Stern, a partner with the law firm WilmerHale based in Washington DC and a senior negotiator under US President Bill Clinton during negotiations on the Kyoto Protocol. “The notion of large reductions from the Chinese perspective seems absolutely undoable.”

War-game organizers say the exercise was intended to engage people with the science and with potential solutions, and to make them test their assumptions in a dynamic situation. From this perspective, they say, the process can be illuminating even if the results are of limited value.

Most participants agree, although Eileen Claussen, president of the Pew Center on Global Climate Change in Arlington, Virginia, says the negotiations, although intense, were not at all representative of what one might expect from India and China in the forthcoming United Nations talks. India would be unlikely to agree so quickly to an emissions-reduction proposal, she says, and China would be unlikely to come to the table empty-handed.

And as a veteran in this arena whose experience in real treaty negotiations dates back to the Montreal Protocol on ozone-depleting gases two decades ago, Claussen says her participation was geared more towards educating others. “My guess is many people learned something from this, but for me it was community service,” she says. ■

Jeff Tollefson



Climate policy expert Eileen Claussen joined scientists and business executives, including Jim Owens (left), chief executive of construction tech firm Caterpillar, in role-playing negotiations.

M. NGAN/AFP/GETTY

Genetic fact-check for ageing story

A genetic study of healthy elderly people aims to uncover the secret of living to a ripe old age — intact.

The 'Welllderly Study' is a joint initiative between the Scripps Research Institute in La Jolla, California, and scientists at the J. Craig Venter Institute in Rockville, Maryland. It hopes to investigate the genomes of 2,000 people aged 80 or more who take no significant medication and have never suffered from any serious disease.

"We are looking at a cohort that we think is harbouring major secrets. They have disease susceptibility genes, but they don't get the diseases you would have expected. Something has protected them. We hope to find out what that is," says study leader Eric Topol, who is director of genomic medicine at Scripps.

Topol and his team will compare gene sequences from their subjects with the same genes in tissues from a control group they've dubbed the 'illlderly'. This second group covers people who died from common, age-related diseases such as cancer, heart attack and stroke before they made it to 80.

Topol and his colleagues Robert Strausberg and Samuel Levy at the Venter Institute

finalized their list of 100 candidate genes last week. It includes genes with an unknown or putative role in healthy ageing, and some that are involved in key jobs such as DNA repair and the handling of insulin. The team plans to expand the list to 500 genes over several years and ultimately to sequence the whole genomes of their elderly recruits. So far, the affiliated Scripps Health System has provided the bulk of the costs of the study.

In principle, the study is "the perfect idea", says Giuseppe Paolisso, a geriatrics expert at the Second University of Naples in Italy, who has studied the relationship between insulin action and longevity. But the problem, he cautions, "is defining which are the major genes".

Topol, whose 360 subjects so far enrolled average 86 years old, is actually working at the youthful end of the spectrum for studies into ageing. Nir Barzilai, for example, who directs the Longevity Genes Project at the Albert Einstein College of Medicine in New York, is

studying Ashkenazi Jews aged 95 and older. Barzilai's team has validated three longevity-related genes so far.

And Rudi Westendorp and Eline Slagboom at Leiden University Medical Centre in the Netherlands have hunted for longevity-associated genetic variations in a group defined by elderly sibling pairs: men aged 89 and above, and women over 91. Results from this study are expected to be published within months.

Westendorp praises Topol's study for "addressing the problem at the heart", by sequencing a wide range of candidate genes. But he says that 80 isn't old enough to rule out a subject surviving in good health simply by chance. "At 80, you are not yet special," Westendorp says. To ensure that genetic variation is involved, "you have to go for the far extreme of the population", he says.

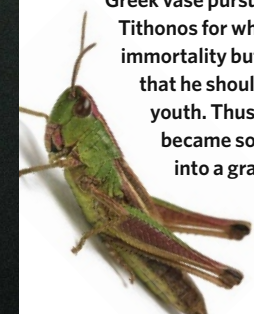
Topol responds that Westendorp's sibling pairs could have had heart attacks, survived three types of cancer and be cognitively impaired. "We have totally intact people, over 80, on no medication. They are distinct outliers," he says. "We have a totally different phenotype — and I think

it's the phenotype that the public is interested in." The search for the genetic roots of longevity is fundamentally different from the welllderly study, says Topol, as his group is focusing on health, not just survival.

Ongoing recruitment to Topol's study is unlikely to be a problem. Subjects "really like to be part of this welllderly group", he says. "It's like a super-senior Facebook. As soon as we recruit people, they start finding all their bridge partners. We are getting calls from all over."

Meredith Wadman

"It's like a super-senior Facebook."



The goddess Eos is depicted on this Ancient Greek vase pursuing her human lover Tithonus for whom she requested immortality but forgot to stipulate that he should also have eternal youth. Thus, he aged until he became so wizened he turned into a grasshopper. As longevity increases worldwide, researchers are concentrating their efforts on achieving a 'youthful' old age.



M.-L. NGUYEN

K. TAYLOR/NATUREPL.COM

**FAITH VALUE**

Did religion arise to protect us from disease?

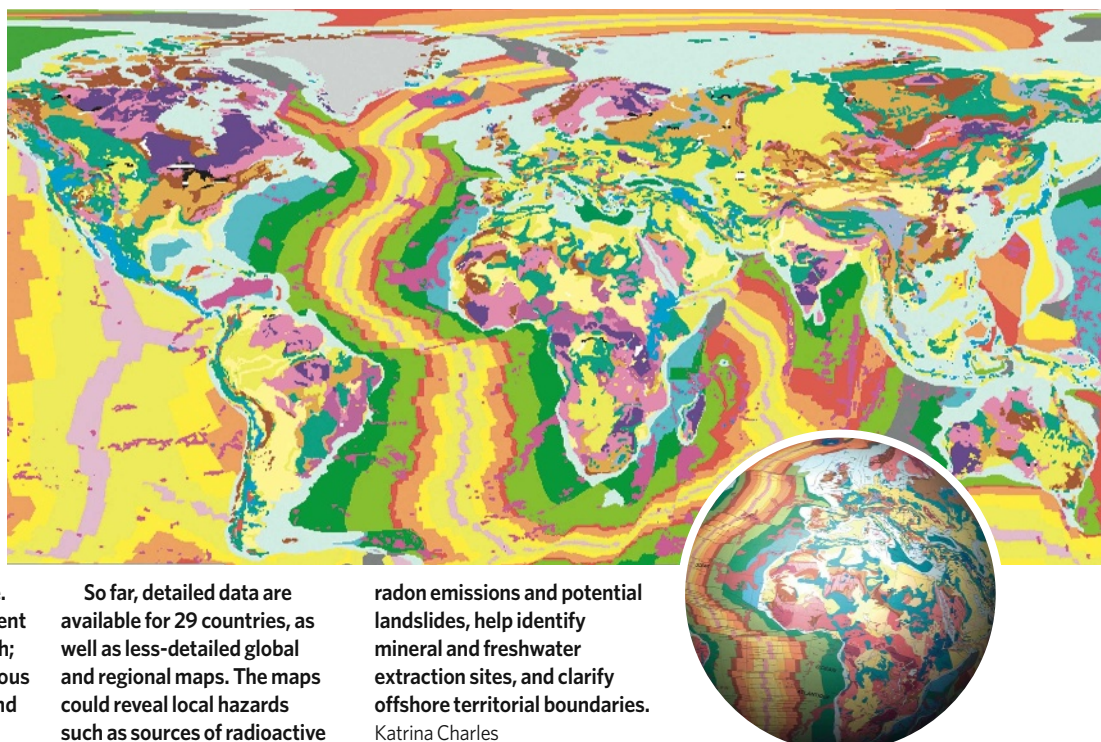
www.nature.com/news

D. BOITEAU/ALAMY
ONEGEOLOGY

SNAPSHOT Earth rocks

The mid-ocean ridge in the Atlantic Ocean, the Rocky and Andes mountain chains along the west of the Americas, the Himalayas in Asia and the Great Rift Valley in Africa can all be seen in the first digital geological map to span the globe.

OneGeology, a global collaboration of computer and Earth scientists, has integrated national and regional geological surveys to produce a free resource that colour-codes the planet by rock age. For example, yellows represent rock from the Miocene epoch; greens are from the Cretaceous period; blues are Jurassic; and pinks are Triassic.



So far, detailed data are available for 29 countries, as well as less-detailed global and regional maps. The maps could reveal local hazards such as sources of radioactive

radon emissions and potential landslides, help identify mineral and freshwater extraction sites, and clarify offshore territorial boundaries.

Katrina Charles

Nerve cells made from elderly patient's skin cells

Skin cells from an elderly patient with amyotrophic lateral sclerosis (ALS) have been 'reprogrammed' to generate motor neurons, the type of nerve cells that die as the disease progresses.

It is the first time that an induced pluripotent stem (iPS) cell line has been created from a patient with a genetic illness (J. T. Dimos *et al. Science* doi:10.1126/science.1158799; 2008). Like embryonic stem cells, iPS cells have the potential to develop into almost any of the body's cell types and offer new disease insights.

The researchers, led by Kevin Eggan of the Harvard Stem Cell Institute in Cambridge, Massachusetts, and Christopher Henderson of Columbia University's Center for Motor Neuron Biology and Disease in New York, made the iPS cells using viral vectors to introduce four genes into skin cells taken from two elderly patients with a mild form of ALS (also known as Lou Gehrig's disease). This genetic reprogramming technique was first developed in 2006 by Shinya Yamanaka of Kyoto University in Japan.

The study shows that iPS cells can be made even from octogenarians with a chronic neurodegenerative disease, although making human iPS cells may be more difficult the older the patient is. Eggan's team was able to generate seven cell lines from the 82-year-old and one

line from her 89-year-old sister — the team went on to characterize just the former's cells.

The paper is expected to be the first in a wave of publications describing the generation of iPS cells from patients with specific diseases. Although results are not yet reported in the peer-reviewed literature, posters at a stem-cell meeting in June described iPS cell lines from people with Alzheimer's disease, Down's syndrome, muscular dystrophy, and more.

Such cell lines could be most useful for diseases that are hardest to research. For example, in ALS, because the dying neurons reside within the spinal cord, they are nearly impossible to study in living patients, says Henderson. "We now have cells in culture that are genetically the same as in those with the disease."

About 2% of ALS patients have mutations in the *SOD1* gene, but in the vast majority the cause is unknown. Mouse models to study the disease depend on limited mutations in this single gene and they do not capture the variations or symptoms seen in individuals. Eggan's team has already collected skin samples from patients with more severe forms of ALS; generating iPS cells from more patients should uncover other genetic and environmental triggers for the disease.

Don Cleveland, an ALS researcher at the

University of California, San Diego, says the results are impressive, but the team still needs to use those iPS-generated cells to "decode the pathways through which mutations cause damage to motor neurons". The researchers must establish "whether immature neurons really reflect the inherited damage that follows only after decades in the human", he says.

The researchers admit that the cells need to be further characterized before they will provide information about disease — Eggan says they should have some answers within months. "We have laid the groundwork for the types of assay that we should do with these cells, and now we're going to do it."

The golden test for clinical use would be to see whether transplanting such cells into patients could alleviate disease — encouraging results have been reported of iPS cell transplants in mouse models of Parkinson's disease and sickle cell anaemia. Cleveland points out that no one has yet been able to create functional motor neurons. And the cell lines created by the team would in any case be unsuitable for therapeutic use in humans, because at least one of the four genes used to program them has a known link to cancer.

Monya Baker

See News Feature, page 682.



BAD VIBRATIONS
Navy report admits sonar's
impact on whales.
www.nature.com/news

PUNCHSTOCK

'Virophage' suggests viruses are alive

The discovery of a giant virus that falls ill through infection by another virus¹ is fuelling the debate about whether viruses are alive.

"There's no doubt this is a living organism," says Jean-Michel Claverie, a virologist at the CNRS UPR laboratories in Marseilles, part of France's basic-research agency. "The fact that it can get sick makes it more alive."

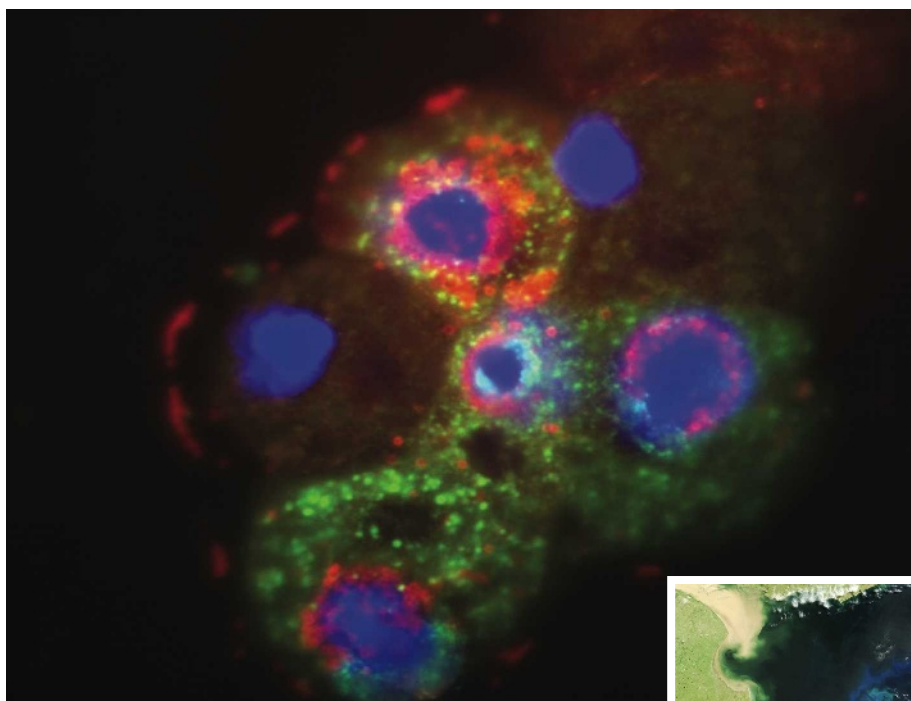
Giant viruses have been captivating virologists since 2003, when a team led by Claverie and Didier Raoult at CNRS UMR, also in Marseilles, reported the discovery of the first monster². The virus had been isolated more than a decade earlier in amoebae from a cooling tower in Bradford, UK, but was initially mistaken for a bacterium because of its size, and was relegated to the freezer.

Closer inspection showed the microbe to be a huge virus with, as later work revealed, a genome harbouring more than 900 protein-coding genes³ — at least three times more than that of the biggest previously known viruses and bigger than that of some bacteria. It was named *Acanthamoeba polyphaga* mimivirus (for mimicking microbe), and is thought to be part of a much larger family. "It was the cause of great excitement in virology," says Eugene Koonin at the National Center for Biotechnology Information in Bethesda, Maryland. "It crossed the imaginary boundary between viruses and cellular organisms."

Now Raoult, Koonin and their colleagues report the isolation of a new strain of giant virus from a cooling tower in Paris, which they have named mamavirus because it seemed slightly larger than mimivirus. Their electron microscopy studies also revealed a second, small virus closely associated with mamavirus that has earned the name Sputnik, after the first man-made satellite.

With just 21 genes, Sputnik is tiny compared with its mama — but insidious. When the giant mamavirus infects an amoeba, it uses its large array of genes to build a 'viral factory', a hub where new viral particles are made. Sputnik infects this viral factory and seems to hijack its machinery in order to replicate. The team found that cells co-infected with Sputnik produce fewer and often deformed mamavirus particles, making the virus less infective. This suggests that Sputnik is effectively a viral parasite that sickens its host — seemingly the first such example.

The team suggests that Sputnik is a 'virophage', much like the bacteriophage



Giant mamavirus particles (red) and satellite viruses of mamavirus called Sputnik (green). Virophages may be common in plankton blooms (inset).

viruses that infect and sicken bacteria. "It infects this factory like a phage infects a bacterium," Koonin says. "It's doing what every parasite can — exploiting its host for its own replication."

Sputnik's genome reveals further insight into its biology. Although 13 of its genes show little similarity to any other known genes, three are closely related to mimivirus and mamavirus genes, perhaps cannibalized by the tiny virus as it packaged up particles sometime in its history. This suggests that the satellite virus could perform horizontal gene transfer between viruses — paralleling the way that bacteriophages ferry genes between bacteria.

The findings may have global implications, according to some virologists. A metagenomic study of ocean water⁴ has revealed an abundance of genetic sequences closely related to giant viruses, leading to a suspicion that they are a common parasite of plankton. These viruses had been missed for many years, Claverie says, because the filters used to remove bacteria screened out giant viruses as well. Raoult's team also found genes related to Sputnik's in an ocean-sampling data set, so this could be

the first of a new, common family of viruses. "It suggests there are other representatives of this viral family out there in the environment," Koonin says.

By regulating the growth and death of plankton, giant viruses — and satellite viruses such as Sputnik — could be having major effects on ocean nutrient cycles and climate. "These viruses could be major players in global systems," says Curtis Suttle, an expert in marine viruses at the University of British Columbia in Vancouver.

"I think ultimately we will find a huge number of novel viruses in the ocean and other places," Suttle says — 70% of viral genes identified in ocean surveys have never been seen before. "It emphasizes how little is known about these organisms — and I use that term deliberately."

Helen Pearson

1. La Scola, B. *et al.* *Nature* doi:10.1038/nature07218 (2008).
2. La Scola, B. *et al.* *Science* **299**, 2033 (2003).
3. Raoult, D. *et al.* *Science* **306**, 1344–1350 (2004).
4. Monier, A., Claverie, J.-M. & Ogata, H. *Genome Biol.* **9**, R106 (2008).

REF. 1

J. SCHWALTZ/NASA

Lawsuit chips away at fish research

Fisheries scientists are continuing an experiment with fish that respond to a dinner bell like Pavlov's dogs, despite a pending US court order that could stop the study.

Food & Water Watch (FWW), a non-profit organization in Washington DC, is suing to halt research by scientists at the Marine Biological Laboratory in Woods Hole, Massachusetts, until further environmental analysis is completed. The laboratory has until 5 September to answer the organization's criticisms, but FWW is confident of a ruling on its request for a temporary injunction before then.

"The purpose of the study is to see if we can selectively harvest hatchery-raised fish after they're released into the ocean," says project leader Scott Lindell. If successful, the research would have implications for restocking wild fish populations and commercial fish farming.

The 5,000 hatchery-raised black sea bass (*Centropristis striata*), which associate a tone with feeding, moved in to their open-sea cage



Will young sea bass come to the call of a bell?

on 17 July. The geodesic 'aquadome' is anchored to the sea floor of Buzzards Bay in Massachusetts. The bass will be released this month and the researchers will then test whether the fish return to the dome when they hear the tone.

The experiment could pave the way for fish ranching in the open sea. "It's like cattle ranching. You don't put up fences, you let them roam around," says Michael Tlusty, director of

research at the New England Aquarium in Boston, Massachusetts. But some say that such a system would have limited use in commercial fisheries. "It is more a biological curiosity than of any relevance to actual wild capture or aquaculture," says Ray Hilborn, a fisheries scientist at the University of Washington in Seattle.

Sound-trained fish could help efforts to boost wild populations, Lindell says, because released hatchery fish would return at the tone for a decent meal while they adjust to open-sea life. "It's to give them sort of an energy packet while they forage," he says.

FWW says that the US Army Corps of Engineers, which issued the permit for the underwater structure, failed to conduct a thorough enough environmental investigation. It is concerned that fish food and waste could pollute the waters around the cage, and that wild fish might learn to associate human activity with food. ■

Amber Dance

S. SENNE/AP

Stanford psychiatrist removed from drug study

Stanford University last week removed its chairman of psychiatry, Alan Schatzberg, as principal investigator from a National Institute of Mental Health (NIMH) grant, after a US Senate investigation raised questions about his impartiality in overseeing the grant.

The grant is for investigating the use of mifepristone, a steroid that is being developed for possible use in depression by Corcept Therapeutics, Menlo Park, California, which Schatzberg helped to found and in which he owns 2.7 million shares worth nearly US\$6 million.

"Despite our belief that Stanford, NIMH and Dr. Schatzberg have handled this in accordance with the regulations and applicable policies and with due regard to the integrity of the research, we can see how having Dr. Schatzberg as the principal investigator on this grant can create an appearance of conflict of interest and we want to eliminate that concern," Debra Zumwalt, Stanford's vice-president and general counsel, wrote in a 31 July letter to the NIMH.

Third failure for private rocket launch

The third attempt to launch the Falcon 1 rocket, built by privately funded company Space Exploration Technologies (SpaceX), has failed. The test flight took off on 2 August from Kwajalein atoll in the Pacific Ocean, but was scuppered minutes after launch when two rocket stages didn't separate.

Falcon 1 was carrying three payloads: a US Department of Defense satellite called Trailblazer, and two NASA satellites, NanoSail-D — which involves a solar sail — and PRESat, a microlaboratory for testing biological samples in space.

SpaceX plans to continue testing Falcon 1 and is developing another rocket, Falcon 9, says chief executive Elon Musk.

Falcon 1's first test flight, in March 2006,



Falcon 1 prepares for its ill-fated launch.

Rock art carves itself a niche online

This series of concentric circles, carved on rocks in Northumberland, UK, is featured in a new catalogue (<http://archaeologydataservice.ac.uk/era>) of carvings from the region.

English Heritage, the government-funded body in charge of protecting historic sites, hopes that three-dimensional models of the carvings uploaded to the Internet will help researchers and those working to preserve the art. The agency says around 2,500 rock carvings are currently known in England, and it expects to discover more with a national survey building on this regional work.



B. KERR/ENGLISH HERITAGE

failed a minute after take-off because of a leaking fuel line. The second, a year later, launched but didn't make it to orbit.

Canadian observatory sold to property developer

The University of Toronto's David Dunlap Observatory has been sold to a development company.

Last week, Metrus Development announced its purchase of the 77-hectare park in Richmond Hill, which includes the observatory, an administrative building and an 1850s farmhouse, for Can\$70 million (US\$68 million). The university will share the proceeds with the Dunlap family, and plans to use its share for a new astronomy and astrophysics institute.

Metrus has pledged to maintain the three buildings, and has invited proposals from anyone interested in reopening and managing the observatory.

The university has run the 1.88-metre reflecting telescope, the largest in Canada, since 1935. But only two faculty members have used it in recent years, along with visiting scientists. "The department of astronomy has gone in a direction for which this observatory is no longer useful," says Robert Steiner, the university's assistant vice-president for strategic communications.

Researcher wins claim for accidental infection

New Zealand's department of labour has admitted that a British researcher who contracted meningococcal septicaemia in 2005 probably became infected while working at a government laboratory.

Jeannette Adu-Bobie, a meningococcal

vaccine researcher, lost both legs, her left arm and the digits of her right hand as a result of the infection. She maintained that she contracted the disease while working at the government's Institute for Environmental Science and Research in Porirua. Previously, the government had said that it was "extremely unlikely" she had contracted the disease there.

But a report released on 3 August by Geraint Emrys, a government adviser, concludes that Adu-Bobie probably did catch the disease at the laboratory. The government has apologized and has offered to pay roughly NZ\$450,000 (US\$330,000) in damages and medical expenses.

Germany tinkers with university regulations

It's not the special law for science that researchers were hoping for, but last week the German government announced a series of regulations to make scientists' lives easier.

Most of German science falls under civil-service rules, which has led to a stultifying bureaucracy at odds with modern needs. But while scientists continue to negotiate for a fundamental change to the constitution, research minister Annette Schavan has identified areas where flexibility can be introduced under existing laws.

For example, as of January 2009, state-funded research institutions and universities will be able to supplement fixed salaries to allow them to attract and keep top scientists. And they will be able to build new labs and other research infrastructure without having to go through lengthy procedures for permission to bypass strict building rules designed for office buildings.

Financial planning

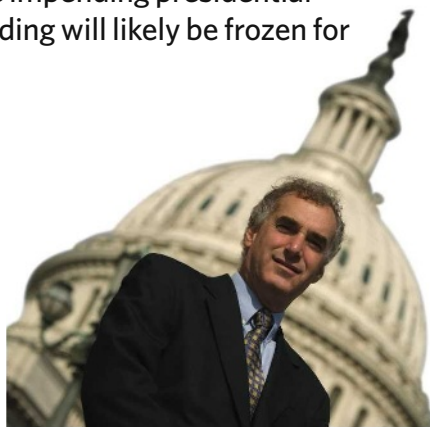
An impasse over the budget and the impending presidential election mean that US science spending will likely be frozen for months, **David Goldston** reports.

Washington DC in August is always a picture of torpor. The humidity rises and Congress escapes for a month-long recess, slowing the pace of work for everyone from Cabinet officials to taxi drivers. But this year, life is even more sluggish than usual, as work on the federal budget has come to a virtual standstill.

Normally, beneath the dormant surface, crucial work on the budget gets done in August. If this were a typical year, agencies would be preparing their spending plans for fiscal year 2010, inaugurating a process that culminates with the president's budget submission to Congress in February. And House and Senate staff would be starting to work out agreements on disputed items for fiscal year 2009, which begins on 1 October. The White House Office of Management and Budget (OMB), the relatively small but powerful unit that oversees budget preparation, would be monitoring and influencing the activities of both the executive and legislative branches. This year, budget staff can pretty much go on holiday like everybody else.

This is not politics as usual. Breaking with precedent, the Bush administration has decided not to prepare a budget for fiscal 2010, given that it will not be in office to either fight for or implement the proposal. In the past, outgoing administrations have left behind a budget as a kind of marker, a last attempt to have an impact on debate. More significantly, the Democratic leaders of Congress have decided to wait until after a new president takes office on 20 January to complete action on the spending bills for 2009, except those for homeland security and defence. As a result, federal agencies, including all the civilian science agencies, will probably have their spending frozen at current levels for at least the first four months of the fiscal year.

At first blush, this is bad news for research funding. Under the current House and Senate versions of the fiscal 2009 spending bills, the major science agencies would receive sizable increases on 1 October. The National Science Foundation, for example, would see its budget jump by about 13% over current levels. But these numbers would probably fall prey to larger budget concerns, as they did last year (see *Nature* 449, 962; 2007), if Congress tried to enact them with President Bush still in office. The Democrats want to spend \$21 billion more



PARTY OF ONE

on domestic programmes than the president has proposed, and Bush has threatened to veto any measure that exceeds his request.

This is the same impasse that emerged at this time last year. The Democrats found out then that they couldn't overcome Bush's determination, despite his sinking popularity. Presidents almost always have the advantage in this kind of test of wills. The legislative process, even at its best, looks messy and confused and involves many voices. The President has easy access to the media and can seem to be above the fray. Moreover, Congress's own popularity ratings have been low. In the end, last year, the Democrats capitulated and accepted the president's overall spending number, dragging science funding down with the total.

The congressional leaders are wise to conclude that the outcome would probably be no different this time. And the Democrats would be just as unlikely this year to find a way to protect science funding if they had to slice \$21 billion from their spending plans without doing violence to their other priorities. So the National Science Foundation and the other agencies might actually be better off with a freeze for the first part of the year, if it means that a big increase is more likely to materialize once a new administration is installed.

Of course, it's not yet clear who the next president will be or what budget he will propose. The Democrats' presumed candidate, Senator Barack Obama of Illinois, has explicitly endorsed doubling funding for basic research (his website doesn't say over what time period), but he has also called for reducing the national debt and offsetting spending increases with cuts elsewhere or with increased revenues. His presumed

opponent, Senator John McCain of Arizona, has called for more spending on biomedical research and has supported science spending as a senator; but he has also called for a one-year spending freeze to figure out how to eliminate the federal deficit by 2013. And whoever wins is likely to face a faltering economy and a record federal deficit that could make it harder to raise spending, even for relatively small accounts like most of the science agencies.

Moreover, the new president won't have much time to figure out how to proceed. Thanks to the stand-off between Bush and Congress, he will not only have to swiftly write his 2010 budget proposal, he will also need to make immediate decisions about fiscal 2009. The system simply wasn't meant to operate this way. But then, compromise was supposed to be the lubricant in the US political mechanism, with its distributed powers. The US Constitution, marvelled a nineteenth-century American notable, set up a "machine that would go of itself". Now it's come to a halt, unable to produce its most basic product, the budget, thanks to ideological polarization, political uncertainty, unwillingness to compromise and personal antipathies.

But behind the scenes, key officials are working to make sure the budget machinery can restart quickly in the next administration. The OMB traditionally puts together budget scenarios and options for the incoming president, and this little-known but vital process is already getting under way. OMB director Jim Nussle, a political appointee, has charged all the federal agencies to submit baseline data this autumn, showing what spending would be needed in 2010 to maintain their current activities. OMB staff, who are civil servants, are working with some agencies on preparing a range of budget options; some agencies, such as the Department of Energy, have decided to put together ideas for three types of budgets, ranging from unrealistically generous to absolutely flat. Once the election is over, the OMB will probably ask the agencies to provide numbers for more options, based, remarkably enough, on what the winning candidate said during the campaign.

So the wheels of government are continuing to turn, albeit slowly and quietly and only at the back end of the machinery. It will be at least February and more likely March before any spending instructions come out from the new president and Congress. And whether that budget will imbue scientists with a mood of summery ease or of wintry constriction is still anybody's guess.

David Goldston, former chief of staff for the House Committee on Science, is a visiting lecturer at Harvard University's Center for the Environment. Reach him at partyofonecolumn@gmail.com.



STANDARD MODEL

Questions raised about the use of 'ALS mice' are prompting a broad reappraisal of the way that drugs are tested in animal models of neurodegenerative disease. **Jim Schnabel** reports.

Several years ago, clinical neurologist Michael Benatar set out to find a drug he could test on some of his patients with amyotrophic lateral sclerosis (ALS). This condition is apt to strike otherwise healthy adults, slowly destroying the neurons that control their muscles. Its roster of famous sufferers includes the physicist Stephen Hawking and the late US baseball star Lou Gehrig.

Except in rare cases, such as Hawking's, ALS progresses inexorably, causing respiratory failure within a few years of diagnosis. Decades of study have revealed a few mutant genes that could cause familial forms of the disease, but no one knows what causes the vast majority of cases — and despite trial after trial of prospective therapies, no therapy has ever been shown to have a major impact on the disease.

One of the first steps Benatar, at Emory University School of Medicine in Atlanta, Georgia, took was to review published data on more than 150 drug tests that other researchers had conducted in the standard mouse model of ALS. Engineered to carry multiple copies of the mutated superoxide dismutase 1 (*SOD1*) gene that causes some inherited ALS cases,

the mouse reliably develops and succumbs to a neuron-killing disease that closely resembles the human condition.

But as Benatar reviewed these mouse studies, he was dismayed to find that the data were of little use to him. Most of the published experiments, including some in top-rank journals, had been done "with small sample sizes, with no randomization of treatment and control groups, and without blinded evaluations of outcomes", he says.

Benatar also found that in the spread of reported results for some drugs, there was statistical evidence that only positive results had been published. Informal conversations with other researchers convinced him that some had tried and failed to confirm reported positive results, but had never published those non-confirmations.

All in all, a body of data that should have yielded useful information was, to Benatar, "questionable at best". He wrote a paper making what he could of the data, and also making clear how flawed he thought they were.

When Benatar's paper¹ was published in *Neurobiology of Disease* in early 2007, it seemed

to have little impact. But in the year and a half since, other investigators have come to similar, and indeed stronger, conclusions. "There is a dawning realization that we may not have designed our mouse drug trials rigorously enough," says Melanie Leitner, chief scientist at Prize4Life, a non-profit organization based in Cambridge, Massachusetts, that promotes ALS research.

That realization is spreading: some researchers are coming to believe that tests in mouse models of other neurodegenerative conditions such as Alzheimer's and Huntington's may have been performed with less than optimal rigour. The problem could in principle apply "to any mouse model study, for any disease", says Karen Duff of Columbia University in New York, who developed a popular Alzheimer's mouse model. In May, a dozen preclinical researchers and mouse model experts thrashed out the issues in a web discussion on the 'Alzforum' website, an online venue where researchers routinely gather to debate neurodegenerative disease issues.

"There has to be sort of a course correction in the field," says Lorenzo Refolo, who oversees grants for preclinical work on neurodegenerative diseases at the National Institute of Neurological Disorders and Stroke in Bethesda, Maryland, "otherwise these practices are just going to continue."

The results of drug tests in mice have never translated perfectly to tests in humans. But in recent years, and especially for neurodegenerative diseases, mouse model results have seemed

ALSTOI

nearly useless. In the past year, for example, three major Alzheimer's drug candidates, Alzhemed (3-amino-1-propanesulphonic acid), Flurizan (tarenflurbil) and bapineuzumab, all of which had seemed powerfully effective in mouse models, have performed weakly or not at all in clinical trials involving thousands of human Alzheimer's patients.

In the case of ALS, close to a dozen different drugs have been reported to prolong lifespan in the SOD1 mouse, yet have subsequently failed to show benefit in ALS patients. In the most recent and spectacular of these failures, the antibiotic minocycline, which had seemed modestly effective in four separate ALS mouse studies since 2002, was found last year to have worsened symptoms in a clinical trial of more than 400 patients².

The minocycline clinical trial, and the other ill-fated ALS-drug trials that preceded it, would never have happened had the prior mouse studies been done properly, says Sean Scott, president of the ALS Therapy Development Institute (ALS TDI), a non-profit biotech company based in Cambridge, Massachusetts. The ALS TDI was set up in 1999 to swiftly screen approved drug compounds for any that could slow the disease in ALS mice. In 2001, it found that ritonavir, an antiretroviral drug used against HIV, seemed to extend lifespan in the mice and the ALS TDI initiated a small safety trial in ALS patients. "But instead of just moving onto the next drug," says Scott, "we kept screening ritonavir in the mice, in the hope that we could learn how it worked and maybe improve on it."

The drugs don't work

Scott and his colleagues were shocked to find that when they scaled up the ritonavir tests by adding more mice, the drug's effect on lifespan didn't become statistically clearer — it disappeared altogether. For the patients in the safety trial, the drug also showed no benefit, and at the highest dose, says Scott, "it was even a little bit detrimental".

Chastened by the experience and with the funding to find out what had gone wrong, the ALS TDI spent the next few years refining its mouse-trial methods and trying to characterize likely sources of error. Its scientists also applied their improved methodological rigour in further tests of prospective ALS drugs, including drugs such as minocycline that had seemed effective in other labs' mouse studies. "But really in the end," says Scott, "we were heartbroken, because even using dramatically more animals than any of those other labs, using very sophisticated pharmacological formulations, and paying attention to levels of the drugs achieved in the mouse nervous systems, we just

could not get any of those drugs to work."

At the time, in late 2006, the ALS TDI was funded in part by a US\$200,000 grant from the Muscular Dystrophy Association (MDA), which has supported ALS research since the 1950s. Sharon Hesterlee, director of research development at the MDA, oversaw the grant and remembers Scott and his colleagues at meetings "taking some flak from members of the academic community, who in effect said 'if you can't reproduce any of our results you're just not doing it right'".

But Hesterlee had already concluded that the ALS TDI "were actually so much more rigorous in their approach, that if there

mouse was not excluded appropriately from a study. Such an exclusion should have happened if, for example, an animal died young of non-ALS causes — Scott came to believe that mouse colonies in some academic labs were "filthy" with infections — or failed to express enough copies of the mutant *SOD1* gene. Further noise could creep in if the treatment and

Design, power, and interpretation of studies in the standard murine model of ALS

SEAN SCOTT¹, JANICE E. KRANZ¹, JEFF COLE¹, JOHN M. LINCECUM¹, KENNETH THOMPSON¹, NANCY KELLY¹, ALAN BOSTROM², JILL THEODOSS¹, BASHAR M. AL-NAKHALA¹, FERNANDO G. VIEIRA¹, JEYANTHI RAMASUBBU¹ & JAMES A. HEYWOOD¹

¹ALS Therapy Development Institute, Cambridge, Massachusetts, and ²Department of Epidemiology and Biostatistics, University of California, San Francisco, USA

Abstract
Identification of *SOD1*^{G93A} as the mutated protein in a significant subset of familial amyotrophic lateral sclerosis (FALS) cases has led to the development of transgenic rodent models of autosomal dominant *SOD1* FALS. Mice carrying 20 copies of the *SOD1*^{G93A} gene are considered the standard model for FALS and ALS therapeutic studies. To date, there have been no published data describing therapeutic agents that extend the lifespan of this mouse. However, no therapeutic agent has been shown to correspondingly extend clinical efficacy. We performed a detailed efficacy study to quantify the



Sean Scott identified sources of noise in mouse model studies for amyotrophic lateral sclerosis.

was anything to find they would have seen it". The MDA soon organized a much larger grant to the ALS TDI, one that would provide it with about \$6 million a year to continue its work. At the same time, Hesterlee also pressured the organization to publish its analysis of the problems with SOD1 mouse tests. "That way our own grant reviewers could start applying those lessons to other mouse studies," she says.

The analysis that the ALS TDI eventually put together overlapped with Benatar's, but also went further: Scott and his colleagues concluded that the previous positive drug trials in ALS mice were likely to have been so plagued by non-drug-related variations in mouse lifespan that this 'noise' was really all they had measured. In the context of small sample sizes and a bias against negative results, they noted, a high degree of noise could easily have led to the appearance of positive results even when no drug effect had existed.

In the ALS TDI's own work, the greatest source of noise had been observed when a

control groups were not evenly matched by gender and litter-membership, because mouse lifespan naturally varied according to these factors. Above all, Scott and his colleagues came to recommend starting each study with at least 24 mice in each treatment or control group — roughly double the norm — to ensure that any real drug effect would statistically rise above the noise.

Scott submitted the ALS TDI findings to the *Proceedings of the National Academy of Sciences* in late 2006 where, he says, "one reviewer flamed us" over the organization's failure to reproduce the positive mouse studies. Scott says he just "wanted to get it out there" and the paper was published in the journal *Amyotrophic Lateral Sclerosis* in January of this year³.

The immediate response was "muted", Leitner observes. "I do think it's a difficult thing for an academic researcher, working as hard and as fast as you can, to have some group tell you that you've made mistakes in your experimental design — especially when that group isn't necessarily operating under the same constraints as you are."

As a biotech company, the ALS TDI had never before published its research in a peer-reviewed journal. Although seven PhD scientists oversaw its laboratory work, its founder, James Heywood, was not a scientist but an engineer who had started the organization after his brother received an ALS diagnosis. Sean Scott had joined the organization as a volunteer after his mother developed the disease. He held a BA in rhetoric from the University of California, Berkeley, and had no formal training as a scientist. And he was now first author on a paper which concluded, in essence, that some high-powered academic researchers had been chasing their tails for years.

Robert Friedlander, at Harvard Medical School's Brigham and Women's Hospital, and the lead author on the first positive study of minocycline in SOD1 mice⁴ defends his work, saying that three other labs independently found similar results. "The fact that ALS TDI did not reproduce these results raises questions as to their methodologies," he says. As for the failed clinical trial of minocycline, Friedlander suggests that the drug may have been given to patients at too high a dose — and a lower dose might well have been effective. "In my mind, that was a flawed study," he says.

Neurologist Jeff Rothstein, who runs a large ALS research lab at Johns Hopkins University School of Medicine in Baltimore, Maryland, says of ALS TDI, "they've done some nice statistics". But the company's failure to reproduce his lab's positive study in 2002 of the anti-inflammatory drug Celebrex (celecoxib) in SOD1 mice⁵ might have been due to differences in study design, he says. Rothstein says that his lab confirmed Celebrex's biological effect at reducing neuroinflammation, whereas ALS TDI didn't look for it. "Were they at variance with us because they never got biological efficacy? Hard to know," says Rothstein. Celebrex later failed in a clinical trial in ALS patients.

The findings of Scott and his colleagues do seem to be resonating now among preclinical researchers. Still, Leitner thinks that too many scientists — including those working on conditions other than ALS — remain unaware that their mouse-model studies may be flawed.

Non-exacting standards

In the recent Alzforum event that Leitner organized, Scott and other researchers discussed the ALS TDI study and in general expressed concerns that these methodological issues went beyond ALS. "People will do an experiment once with ten animals and get a result, and if it's the right result it gets published in a high-profile journal," says Duff, who has followed the ALS TDI story with keen interest. "And there's no requirement that you show the effect again with a different set of mice, or in a larger group of mice, or in a different model."

"Many of the neurodegenerative models out there are on mixed and segregating genetic backgrounds," says Greg Cox, a forum participant and mouse geneticist who works for



"There is a dawning realization that we may not have designed our mouse drug trials rigorously enough."
— Melanie Leitner

the world's largest provider of research-grade mice, the Jackson Laboratory in Bar Harbor, Maine. "So in those cases even [inbred] littermates aren't genetically identical. And if a genetic background effect [for example, on mouse lifespan] is there, you could end up measuring those background effects more than the drug effect you want to focus on."

Mike Sasner, a neuroscientist also at the Jackson Laboratory, notes that spontaneous genetic changes often affect the disease-causing mutant gene directly. For SOD1 mice, and for Alzheimer's mice engineered to overexpress amyloid protein, changes are often seen in the number of copies of

the disease-driving 'transgene'. For Huntington's disease mice, spontaneous changes can alter the number of disease-causing repeat sequences within the transgene. "So if you're breeding these Huntington's mice over ten generations and you go from 100 repeats to 50 repeats, you're going to basically lose the phenotype," says Sasner.

He adds that the Jackson Lab now checks the genetic make-up of its transgenic mice to reduce the problem, but not every researcher is aware of such issues. "You might create a mouse in your lab and distribute it to ten different people," he says, "so there's ten different colonies all over the world, and they're all diverging from each other, genetically. So when I'm publishing my paper I'm talking about this mouse, and you think you have the same

mouse in your hands. But do you really?"

Sasner says that he and other scientists now hope to draft a formal document with guidance for preclinical researchers on these issues. Leitner applauds such efforts, but thinks the National Institutes of Health (NIH) needs to get involved too. "I believe that if the government doesn't encourage consideration of these issues and support them, it's going to be a very hard sell," she says. "Because basically the ALS TDI study suggests everyone needs to conduct much more thorough and expensive animal-model trials."

Duff agrees. "There just aren't the resources now to do really large, well-powered mouse studies. So I think the NIH should get ahead of the curve here, for example setting up a programme and sending out Requests for Applications to study this, to look more at the mouse models and, in the same way that Sean Scott's group did, to see why they're failing to translate."

Even Refolo thinks that the agency should take the lead in addressing these problems. "I think there just has to be a [new] policy," he says. But he emphasizes that the NIH grant process is regulated largely by the academic community itself. "The academics have to embrace the policy, and people who are reviewing grants and papers have to embrace it and also the editors of these journals have to embrace it." And so far Refolo sees no sign of a major change coming. "If [these issues] are being addressed in study section, where these grants are reviewed, it's at least below my radar."

Even if preclinical researchers ultimately do clear up the methodological flaws in mouse studies, they'll have other issues to deal with. For example, as Benatar noted in his 2007



Model behaviour: macaques have been engineered to carry the human Huntington's disease gene.

M. METZ

YANG, S.-H. ET AL NATURE 453, 921-924 (2008)



Non-profit biotech ALS TDI challenged some of the methods commonly used in drug tests on mice.

paper, SOD1 mice are typically treated with drugs well before the onset of symptoms. Yet in nearly all human cases, Benatar says, “we have no capacity to initiate therapy presymptomatically, so it’s wishful thinking to suppose that success with presymptomatic treatment in a mouse is going to translate into efficacy in a human”.

The wrong model?

Perhaps the biggest issue facing the field is whether the mouse models faithfully reproduce the biology of the human disease. Alzheimer’s mouse models typically develop amyloid ‘plaques’ in their brains, but they do not develop an Alzheimer’s-like dementia and anti-amyloid strategies have repeatedly failed to slow the disease in clinical trials. Parkinson’s researchers have never had a good mouse model for the full disease process, and even the mouse model for Huntington’s disease — a relatively simple genetic disease — does not fully reproduce the clinical signs seen in humans with the disorder.

SOD1 mice have often been considered one of the most accurate animal models of any neurodegenerative disease. Although the SOD1 gene is mutated in only around 20% of human familial cases (representing 2–3% of all ALS) the disease that the animals get so closely resembles the common, sporadic form of human ALS

that the two maladies have been assumed to share a ‘final common pathway’ of neuronal destruction.

That assumption is now being questioned. In the past two years, researchers have found evidence suggesting that a DNA-binding protein, TDP-43, could be the trigger for sporadic ALS when it is malformed or improperly processed inside cells — and that SOD1-driven ALS might really be a distinct disorder. “The idea that the SOD1 model could be extended to sporadic ALS patients is not holding up,” says Virginia Lee, a neuropathologist at the University of Pennsylvania in Philadelphia, whose lab was the first to report the link between TDP-43 and sporadic ALS⁶. Her lab is one of several now racing to devise a mouse model of ALS carrying mutated TDP-43.



“I’m talking about this mouse, and you think you have the same mouse in your hands. But do you really?”

— Mike Sasner

The debate over SOD1’s relevance to the majority of human ALS cases feeds a broader worry, namely that it may be unrealistic to think of modelling the full complexity of ageing-related human brain disorders in mice whose disease course is usually accelerated by a single, crude genetic modification.

In theory, with their more human-like nervous systems, monkeys should make much better models of human neurodegenerative diseases. Aged vervets have already been used to test Alzheimer’s vaccines; and Anthony Chan and his colleagues at the Yerkes National

Primate Research Center in Atlanta, Georgia, have described the creation of macaques that carry the human Huntington’s disease gene, and suffer from a very similar disorder⁷. “I definitely think that the non-human primate models need to be brought to the forefront much more than they have been,” says John Morrison, an animal-model researcher at Mount Sinai School of Medicine, New York.

But aside from the moral issues that this would raise for some researchers, a switch to monkeys could be prohibitively expensive. Scott estimates that a six-month study of about 50 ALS mice would cost roughly \$100,000, whereas Stuart Zola, who heads the Yerkes Center, estimates that more than \$500,000 would be needed to study 50 ordinary macaques for two years — and possibly much more time and money would be required to study drugs for slow-burning neurodegenerative diseases whose effects may only become apparent with advanced age.

Scott, Duff, Rothstein and others suggest that mouse models should still be used, but that drug tests in them should target specific, disease-related molecular pathways — for example, TDP-43 accumulations, if they turn out to be relevant — instead of broader endpoints such as lifespan or behaviour, where mice and men are inherently mismatched. Scott says he now sees the SOD1 mouse as “perhaps a pathway model as opposed to a disease model, and if we can affect survival [with drugs], great”. The ALS TDI is currently finishing a study of gene expression patterns in various tissues from SOD1 mice and humans with ALS, to find what molecular pathways of disease they share — if any.

Mouse models could therefore end up being not only more difficult and expensive to use with acceptable rigour, but at the same time more narrowly predictive of the human condition. But whether preclinical researchers will accept such a radical change remains to be seen. “I think there’s a sense of desperation that we need a convenient model for bringing drugs to clinical trial,” says Benatar. “And I do sort of hear that concern.” But desperation, he adds, is an inadequate justification for the continued use of a poor model. “It’s a bit like the proverbial drunk who keeps looking for his lost keys under the lamp post, simply because the light’s better there.” ■

Jim Schnabel is a freelance writer based in Maryland.

1. Benatar, M. *Neurobiol. Dis.* **26**, 1–13 (2007).
2. Gordon, P. H. et al. *Lancet Neurol.* **6**, 1045–1053 (2007).
3. Scott, S. et al. *Amyotroph. Lateral Scler.* **9**, 4–15 (2008).
4. Zhu, S. et al. *Nature* **417**, 74–78 (2002).
5. Drachman, D. B. et al. *Ann. Neurol.* **52**, 771–778 (2002).
6. Neumann, M. et al. *Science* **314**, 130–133 (2006).
7. Yang, S.-H. et al. *Nature* **453**, 921–924 (2008).

ALS TDI

JACKSON LABORATORY



HARVARD UNDER REVIEW

Harvard is embarking on an experiment to foster collaboration and interdisciplinary research. **Corie Lok** looks at whether it can change its culture and reinvent communities along the way.

Harvard University is the oldest, richest, most prestigious institution of higher learning in the United States. In many fields it is also the best. But its efforts to move into new areas of interdisciplinary science are often slow or fragmented compared with more nimble powerhouses such as Stanford University or the University of California, Berkeley.

A typical example is chemical biology, in which researchers use small molecules to probe and learn about biological systems. Harvard has one chemical-biology group on its main campus in Cambridge, Massachusetts; another at the Harvard Medical School, six kilometres away across the Charles River in Boston; yet another at the Broad Institute, jointly run by the Massachusetts Institute of Technology and Harvard; and still other, smaller efforts within some of Harvard's many affiliated research hospitals and institutions.

Such a proliferation of similar efforts troubles

Tim Mitchison, co-founder of Harvard Medical School's Institute of Chemistry and Cell Biology. "It leads to inefficient use of resources and expertise," he says.

It troubles Harvard's leadership, too — which is why the university is now engaged in a controversial effort to reinvent the way that research gets done and managed there.

A major part of this effort is to expand the campus in Allston, a neighbourhood of Boston just across the river from Cambridge. Construction began there earlier this year on a 46,500-square-metre, US\$1-billion facility that will house the new Department of Stem Cell and Regenerative Biology. The department was formed last year as a new academic home for Harvard's stem-cell community, which, like the chemical-biology groups, is currently scattered across various schools, research hospitals and centres. Bioengineering and systems-biology researchers will also be moving to the new complex.

Another key part of the effort is the Harvard University Science and Engineering Committee (HUSEC), which encompasses Harvard's provost and other top leaders of science and aims to improve the coordination of science planning across the university and to provide funding to start up new initiatives.

Yet another important piece is a strategic review now under way at the medical school — the first in recent memory. Instigated by the school's new dean, Jeffrey Flier, the review is asking for recommendations on how to strengthen ties across the medical school and research hospitals in areas such as human genetics, neuroscience, technology development and therapeutics and chemical biology. A similar review is under way at the Faculty of Arts and Sciences, led by its new dean, Michael Smith.

In effect, Harvard has launched a radical, university-wide experiment to create a more open, agile and collaborative research culture — as well as research communities that go

BEHNISCH ARCHITECTEN/ALLSTON DEVELOPMENT GROUP, HARVARD UNIV.

Harvard is building a new science complex at its Allston campus to boost interdisciplinary science.

beyond traditional departments and schools. "Science is changing in dramatic ways," says Harvard's president Drew Faust. "The implications of this for the way we organize our scientific enterprise are very significant. It challenges us to break down barriers." Like the president, the provost and deans at Harvard have also bought into this vision. "We're realizing that we can be more than the sum of our parts," says Jeremy Bloxham, dean of science in the Faculty of Arts and Sciences. But what remains to be seen is whether enough faculty members and departments at a grassroots level will join them for the changes to take hold.

Grand designs

Harvard has the resources to show on a grand scale how to foster new ways of doing science. It has a large and growing science and engineering component — 1,750 investigators, including those from Harvard's 20 affiliated hospitals and research institutes — which pulls in around \$2 billion a year in external funding. At \$35 billion, it has the biggest endowment of any university, and it has more than 80 hectares of land on which to expand its Allston campus.

But Harvard is used to being a leader and isn't in the habit of self-reflection or admitting to its weaknesses. It has a long tradition of fierce independence — every school and hospital has its own powerful leaders, administrative structures, fundraising efforts and cultures. At times, hiring committees from different parts of Harvard even compete with each other for the same candidate. And researchers doing similar work may be more likely to come across each other at international meetings than at home. Until recently, says Venkatesh Narayanamurti, dean of Harvard's School of Engineering and Applied Sciences, "our organizational structure was always very decentralized. The president had very little authority, and the money was with the deans. That's not all bad, because you want decisions being made close to where the work is being done. But one has to balance the role of local autonomy with central direction."

Bringing more centralized direction to Harvard was a key goal of Larry Summers, the university's president from 2001 to 2006. To help him implement that strategy, he brought in Steve Hyman, a psychiatrist at Harvard Medical School from 1992 until 1996 and then director of the National Institute of Mental Health. Summers installed Hyman as provost

in 2001 and together the pair moved aggressively to change the status quo. By the end of 2003, Harvard Medical School had established a new systems-biology department; the university had announced it would collaborate with the Massachusetts Institute of Technology to launch the Broad Institute; and Summers had convened four task forces to jump-start planning of the Allston campus.

One of those task forces, chaired by Hyman and focused on science and technology, envisaged that Allston would become a hub for interdisciplinary science. Accordingly, it put out a call for ideas for science projects that needed space and investment. From the 70 proposals it received, several interdisciplinary, cross-school initiatives were selected to receive start-up funding from the provost's office, including the Initiative in Innovative Computing, the Microbial Sciences Initiative and the Origins of Life Initiative.

Meanwhile, the university was drawing up ambitious plans for a large fundraising campaign, with substantial sums devoted to science. University spokesman John Longbrake explains that Harvard generally needs to raise funds to support new initiatives, as 83% of the endowment is restricted by donors for specific uses, such as undergraduate financial aid or library collections, and it spends only about 5% of its entire endowment every year.

Then trouble hit: in early 2005, Summers made his infamous remarks about how innate differences between men and women affect their success in science. The resulting headlines galvanized long-simmering discontent with Summers, whose aggressive, top-down leadership style had proved highly polarizing. Although some faculty members found his vision and decisiveness refreshing and exciting, others felt disenfranchised and offended.

Fast and furious

Many critics also thought that the Allston task-force planning process was moving too quickly, and wasn't inclusive enough; and some even questioned the need to expand so aggressively into Allston at all. Others considered the creation and funding of the new interdisciplinary initiatives to be too ad hoc, and that more transparent governance structures and consultation were needed. "There was a bit of a free-for-all period," says James McCarthy, a biological oceanographer at Harvard. "There needed to be some order."

In 2006, with so much unrest on campus, Summers was forced to resign after just five years in the job (see *Nature* 10.1038/news060220-13; 2006). As one of his last acts as president, he convened a university-wide faculty committee to address concerns about how initiatives that involve researchers from different departments and schools should be organized and managed. In its final report at the end of 2006, the committee recommended the formation of the HUSEC and other organizational structures for science, such as cross-school departments and

"We're realizing that we can be more than the sum of our parts."

— Jeremy Bloxham



Harvard's president, Drew Faust, has to win over students and staff alike.

J. IDE/HARVARD UNIV. NEWS OFFICE

interdepartmental committees. The 17-member HUSEC, chaired by Hyman, began meeting in April 2007. Among other things, its \$50-million annual budget now funds the initiatives launched as a result of the Allston task forces.

The scientists involved with those initiatives are perhaps the ones who have been the most affected by Harvard's bumpy foray into interdisciplinary science. They began with much enthusiasm during Summers's reign, launching new research projects, holding symposia and seminars, recruiting graduate students, postdocs and other staff, and churning out papers. But with Summers's departure, a stalled fundraising campaign and shifts in their governance structure, the initiatives have lost momentum. Some have not received the funding they had initially planned for, and they have not been able to take on new research projects.

Fighting the system

Harvard's experiment is as much about leadership as it is about science. Summers got the process moving, but in a way that also generated much opposition. Now the job falls to his successor, Faust.

Faust is not a newcomer to leading institutional change. As the first dean of Harvard's Radcliffe Institute for Advanced Study, where she served for six years, Faust oversaw the transformation of the former independent women's college into a Harvard-affiliated scholarly institute. Initially, sceptics questioned whether this was the right type of experience for the Harvard presidency; the Radcliffe Institute is small and doesn't have a faculty or students, just fellows.

But so far, Harvard scientists have generally been pleased with Faust, saying that she's insightful, respectful, supportive of faculty and a good listener. Although a historian, Faust has been supportive of science, testifying before the US Senate in March calling for greater funding from the National Institutes of Health. She has kept Hyman on as provost, effectively allowing science planning to continue without major changes in direction.

Faust has put unity high on her agenda, speaking of Harvard as "one university" at which collaborative research can flourish, rather than just a loose collection of independent schools. And she has hired new deans who share that vision. Smith, for instance,

says that "one of the things we want to do here is not just push forward with individual basic research in the sciences and engineering, but also have connections made with the professional schools."

The next test of Faust's leadership will come with the revival of the fundraising campaign. As fundraising is so crucial to Harvard's ability to launch new projects, Faust's challenge will be to win over potential donors.

Over the next several years, Harvard will face

In fact, the departments within Harvard's Faculty of Arts and Sciences are beginning to loosen up. They are doing more joint appointments. In some tenure cases, letters of recommendation from people outside of the candidate's home department are given equal consideration. And the medical school has revised its promotion criteria for its faculty to explicitly include collaboration with groups of investigators.

Another major issue is how to fund interdisciplinary projects. Alyssa Goodman, founding director of the Initiative in Innovative Computing says, "I would say that half of the mission of our initiative is to do interdisciplinary science and the other half is to come up with a rational funding model for how you should do that."

Dimitar Sasselov, director of the Origins of Life Initiative also thinks a lot about funding. Because his initiative doesn't fit neatly into the traditional categories of research funded by government agencies, it is especially dependent on seed money from the university and on private sources during the first few years. And that money needs to come quickly, says Sasselov, while the inspiration is still fresh. Otherwise, busy collaborators are likely to simply drift apart. What is required, he says, is a mechanism that would support interdisciplinary projects in the way that a venture-capital firm invests in a start-up company: quick but limited funding, with clear timelines and milestones.

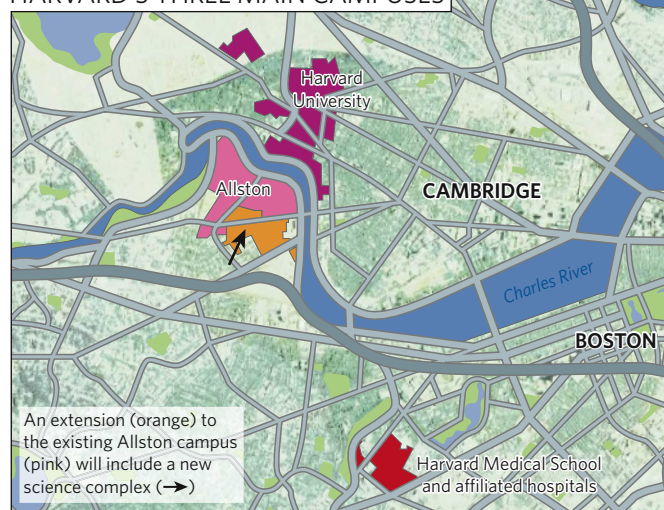
Business approach

Business approach

That 'venture-capitalist' function is one of the multiple roles that Hyman sees for the HUSEC. So far, the committee has put out a call for proposals for seed grants of up to \$75,000 per year for interdisciplinary research. It has also designated \$7 million of its annual budget for the funding of first-year science graduate students. In addition, it has commissioned a review of how Harvard can beef up its bioengineering endeavours. And it has come up with a plan for what will go into the first science building on the Allston campus.

Despite those steps forward, the HUSEC remains a work in progress. The committee has yet to develop a clear mission and to decide how it will prioritize its projects — not to mention put in place a system for allocating its funds. A particular challenge, many say, is to strike the right balance between support for the existing, core disciplines, and investments in the new ones.

HARVARD'S THREE MAIN CAMPUSES



"Harvard has an opportunity to play a leadership role in the world in solving some of the great challenges like climate change."

— Dan Schrag

many questions as it expands and reconfigures science. "There's a bit of naivety in terms of how difficult it is to get these things to work, taking it beyond just talk," says Geoffrey West, president of the Santa Fe Institute in New Mexico, which specializes in interdisciplinary research. "It takes much more than just setting up [a new centre or building]. Universities need to provide mechanisms for this kind of research to be rewarded in the right ways and for people to interact both informally and formally." In particular, he says, they have to adapt their hiring, tenure and promotion criteria to make them more suitable for young interdisciplinary researchers who work with people from fields outside their own department.

This will be a major challenge, given how deeply entrenched department-based structures and cultures are at many universities, says West. "You've got to get people [from different departments and fields] to break away from their departments and spend serious time together," says West. "I'm not an anti-department guy. I like departments. They just have to loosen up."



Harvard's main campus is the historic centre of the university.

Perhaps the most prominent test case is the Department of Stem Cell and Regenerative Biology. Chaired by stem-cell researchers Doug Melton and David Scadden, it will be the first tenant in the Allston science building. It developed out of the Harvard Stem Cell Institute, also co-directed by Melton and Scadden, as a way to enhance undergraduate teaching and to hire and promote new faculty members, as institutes do not have the authority to make these decisions.

Although its new building won't open until 2011, the department is already active. Several existing faculty members have moved to the new department and some will begin teaching undergraduates this year.

Moving plans

However, says Scadden, creating the department is one thing. It will be much more complicated to figure out which members of the department will move their labs to the Allston campus, and whether they'll move over completely or maintain labs in two locations. Many stem-cell researchers who are scattered across different campuses and hospitals are excited about the prospect of being next-door neighbours in Allston. But some are concerned about losing connection with their existing communities. For instance, Gary Gilliland, head of the

Cancer Stem Cell Program at the Harvard Stem Cell Institute and based in the medical campus says he is interested in Allston, but that "it's possible I would have a harder time recruiting clinical fellows" there.

Home institutions and departments are also concerned about losing key faculty members to the Allston campus and having their communities disrupted. "It's good for them and it's maybe not good for us," says neuroscientist Jeff Lichtman, referring to the three people from the molecular and cellular biology department in which he works, including Melton, who will be moving to the new stem-cell department. "We're not losing them as colleagues; some will have joint appointments. But the truth is distance is a big problem for scientists."

"We want to make sure that we do everything we can to not gut those communities," says Scadden. "We want them to be vibrant. It's essential to the success of this department and for the Stem Cell Institute that we continue to have close relationships with the hospitals and the medical school."

Some of the junior faculty members are deferring their decision until they see which

of the senior researchers move to Allston. "It's important for us to know who will be going there, who our neighbours would be," says Konrad Hochedlinger, a stem-cell biologist at the Harvard Stem Cell Institute, who will be part of the new department.

As the university builds Allston and plans for the next science buildings there, discussions will continue about which other scientific communities will be the most appropriate ones to move there. For some emerging fields such as systems biology or regenerative biology, creating a critical mass in one location may be the key but for other, more mature fields, it may be more important to have individuals scattered throughout an established campus where they can easily interact with people from other departments and disciplines.

"I like the model of distributed expertise as opposed to critical masses," says Lichtman, who takes advantage of his close proximity to the physical-sciences departments to attend seminars and meet colleagues there. "It's not that I know everything a neuroscientist might say to me, but I would know a lot of what a neuroscientist might say to me because I've been talking to neuroscientists most of my career. But every time I talk to a chemist or a physicist, I learn something new."

Testing times

In the end, perhaps the most important challenge facing Faust during this era of change is to find the right combination of top-down decision-making and bottom-up consensus-building — especially given how much her predecessor failed at this.

Some faculty members who liked Summers's approach hope that Faust, who is seen as being more considered in her leadership style, won't be so cautious that Harvard settles back into its old, slow ways. "Harvard has an opportunity to play a leader-

ship role in the world in solving some of the great challenges such as climate change and I really hope that, over the coming year or two, Faust does stand up and really lead on these issues," says Dan Schrag, director of the Harvard University Center for the Environment.

Faust is taking her own approach to leadership. "We need to be careful about what we do but we also need to figure out what the important initiatives are and then do them," says Faust. "I don't decide to do things because they're bold. I decide to do things because they're right."

Corie Lok is senior editor of Nature Network, based in Boston.

"Universities need to provide mechanisms for people to interact both informally and formally."

— Geoffrey West

CORRESPONDENCE

Small countries are unexpected winners in ERC grant tables

SIR — Your Editorial ‘Supporting the future’ (*Nature* **453**, 958; 2008) and News Feature ‘The research revolution’ (*Nature* **453**, 975–976; 2008) rightly point out that the European Research Council (ERC) will change the future of European science funding for the better. But the conclusion that most of the ERC’s young-investigator research grants go to the top three of European research — the United Kingdom, France and Germany — disguises a less impressive performance by these countries when the numbers of research grants awarded are considered relative to the overall population or gross domestic product (GDP) (see table).

It seems that the countries where the greatest number of grants are obtained per capita — and hence with the most efficient young scientists — are Cyprus, Israel, Switzerland and

the Netherlands. Bulgaria, which won just one grant, does better than the traditional strongholds in terms of grants in relation to GDP, and Hungary with six is in third position.

Germany, along with Portugal, Norway and the Czech Republic, does poorly, trailing Italy and Greece in grants per GDP. This might be partly because Germany doles out plenty of its own young-investigator grants, which are worth as much as the highly competitive ERC grants. Young German researchers, unlike their Italian counterparts, for example, might therefore not feel so desperate to pursue European money.

What other surprises may come to light, after further analysis and contextualization of the useful numbers provided by this first competition for ERC young-investigator grants? **Friedrich Frischknecht** Department of Parasitology, University of Heidelberg Medical School, INF 324, 69120 Heidelberg, Germany **e-mail: freddy.frischknecht@med.uni-heidelberg.de**

Country	Grants*	Population (millions)†	GDP (US\$ bn)†	GDP per capita (US\$)	Ranking, grants per capita	Ranking, grants per overall GDP
UK	58	59.7	2,199	36,830	7	10
France	39	60.5	2,127	35,150	9	12
Germany	33	82.7	2,795	33,800	16	16
Netherlands	26	16.3	624	38,290	4	5
Italy	26	58.1	1,763	30,340	15	14
Israel	24	6.7	123	18,420	1	1
Spain	24	43.1	1,125	26,090	11	11
Switzerland	16	7.3	367	50,280	3	4
Sweden	11	9	358	39,740	6	8
Belgium	10	10.4	371	35,660	8	9
Finland	7	5.2	193	37,000	5	7
Hungary	6	10.1	109	10,820	10	3
Greece	4	11.1	225	20,290	17	13
Austria	4	8.2	306	37,330	13	15
Denmark	3	5.4	259	47,910	12	17
Cyprus	2	0.7	33	46,870	2	2
Ireland	2	4.1	202	49,220	14	19
Portugal	2	10.5	183	17,460	19	18
Norway	1	4.6	296	64,240	18	21
Bulgaria	1	7.7	26.6	3,460	20	6
Czech Rep.	1	10.2	124	12,190	21	20

*ERC figures from *Nature* **453**, 976 (2008).

†From *The Economist Pocket World in Figures* (2008); data on Cyprus are from the Internet.

Human microbiome: hype or false modesty?

SIR — Your Editorial ‘Who are we?’ (*Nature* **453**, 563; 2008) warns that research into the human microbiome should not be oversold until medical promise is established. Fads, hype and false promises have no place in any research, but neither should we bias debate by suppressing reasonable extrapolation.

We were led to believe, wrongly, that genome sequences would open our eyes and we would understand — but that did not reduce the value of those sequences in the slightest. They are more valuable than expected precisely because we did not understand them as expected. Research showing us that we do not comprehend what seems to be obvious can be important. Unfortunately, it is the most difficult to conceive, justify, fund and publish because, until we do, we think its premises are wrong.

Decisions on whether to fund particular projects are not scientific decisions. They are social decisions, and may therefore be over-influenced by hype. Research funding must compete with other funding needs, and research hype is up against the hype of military spending, spiralling oil costs and the latest celebrity exploits. Waiting for medical benefits to be proven before funding basic research would mean that basic research would never be funded.

Medicine is founded on germ theory, physiology, hygiene and antibiotics. Yet there are profound gaps in our understanding of those interactions. Health conditions common in the developed world (obesity, diabetes, allergies, asthma, heart, vascular and inflammatory diseases) remain rare to non-existent in the rural undeveloped world, despite the lack of

potential treatment. The ‘hygiene hypothesis’ tries to explain these differences, but so far no protective agent has been found. If the microbiome project can eventually deliver such an agent, or agents, and stop even a fraction of the health decline associated with economic development, the effects will still be profound.

The analogy of the microbiome with rainforest biodiversity is apt. We should exhaustively sample and understand the diversity of both in the wild before the ‘clear cutting’ of modern agriculture and hygiene practices irreversibly destroys them both.

David R. Whitlock Nitroceutic LLC, Watertown, Massachusetts 02472, USA

e-mail: daedalus4u@yahoo.com

Competing financial interests: D. R. W. has patented the use of autotrophic ammonia-oxidizing bacteria as a topical treatment.

Human microbiome: take-home lesson on growth promoters?

SIR — Your News Feature ‘The inside story’ (*Nature* **453**, 578–580; 2008) highlights the potential importance of the gut microbiome in understanding health. As you discuss, human profiles differ in obese and lean individuals and are being scrutinized for their possible influence on weight.

A factor you don’t mention is the routine administration over some 50 years of antibiotics and probiotics as growth promoters to farm animals. Use of antibiotics for this purpose was banned in 2005 in the European Union, but it continues in many places (including the United States).

Compounds with antibiotic activity and bacteria with probiotic activity have been widely tested as growth promoters (see, for example, M. Khan *et al. Brit. Poultry Sci.* **48**, 732–735; 2007). The probiotics used in agricultural industries are mainly *Lactobacillus*

CORRESPONDENCE

Small countries are unexpected winners in ERC grant tables

SIR — Your Editorial ‘Supporting the future’ (*Nature* **453**, 958; 2008) and News Feature ‘The research revolution’ (*Nature* **453**, 975–976; 2008) rightly point out that the European Research Council (ERC) will change the future of European science funding for the better. But the conclusion that most of the ERC’s young-investigator research grants go to the top three of European research — the United Kingdom, France and Germany — disguises a less impressive performance by these countries when the numbers of research grants awarded are considered relative to the overall population or gross domestic product (GDP) (see table).

It seems that the countries where the greatest number of grants are obtained per capita — and hence with the most efficient young scientists — are Cyprus, Israel, Switzerland and

the Netherlands. Bulgaria, which won just one grant, does better than the traditional strongholds in terms of grants in relation to GDP, and Hungary with six is in third position.

Germany, along with Portugal, Norway and the Czech Republic, does poorly, trailing Italy and Greece in grants per GDP. This might be partly because Germany doles out plenty of its own young-investigator grants, which are worth as much as the highly competitive ERC grants. Young German researchers, unlike their Italian counterparts, for example, might therefore not feel so desperate to pursue European money.

What other surprises may come to light, after further analysis and contextualization of the useful numbers provided by this first competition for ERC young-investigator grants? **Friedrich Frischknecht** Department of Parasitology, University of Heidelberg Medical School, INF 324, 69120 Heidelberg, Germany **e-mail: freddy.frischknecht@med.uni-heidelberg.de**

Country	Grants*	Population (millions)†	GDP (US\$ bn)†	GDP per capita (US\$)	Ranking, grants per capita	Ranking, grants per overall GDP
UK	58	59.7	2,199	36,830	7	10
France	39	60.5	2,127	35,150	9	12
Germany	33	82.7	2,795	33,800	16	16
Netherlands	26	16.3	624	38,290	4	5
Italy	26	58.1	1,763	30,340	15	14
Israel	24	6.7	123	18,420	1	1
Spain	24	43.1	1,125	26,090	11	11
Switzerland	16	7.3	367	50,280	3	4
Sweden	11	9	358	39,740	6	8
Belgium	10	10.4	371	35,660	8	9
Finland	7	5.2	193	37,000	5	7
Hungary	6	10.1	109	10,820	10	3
Greece	4	11.1	225	20,290	17	13
Austria	4	8.2	306	37,330	13	15
Denmark	3	5.4	259	47,910	12	17
Cyprus	2	0.7	33	46,870	2	2
Ireland	2	4.1	202	49,220	14	19
Portugal	2	10.5	183	17,460	19	18
Norway	1	4.6	296	64,240	18	21
Bulgaria	1	7.7	26.6	3,460	20	6
Czech Rep.	1	10.2	124	12,190	21	20

*ERC figures from *Nature* **453**, 976 (2008).

†From *The Economist Pocket World in Figures* (2008); data on Cyprus are from the Internet.

Human microbiome: hype or false modesty?

SIR — Your Editorial ‘Who are we?’ (*Nature* **453**, 563; 2008) warns that research into the human microbiome should not be oversold until medical promise is established. Fads, hype and false promises have no place in any research, but neither should we bias debate by suppressing reasonable extrapolation.

We were led to believe, wrongly, that genome sequences would open our eyes and we would understand — but that did not reduce the value of those sequences in the slightest. They are more valuable than expected precisely because we did not understand them as expected. Research showing us that we do not comprehend what seems to be obvious can be important. Unfortunately, it is the most difficult to conceive, justify, fund and publish because, until we do, we think its premises are wrong.

Decisions on whether to fund particular projects are not scientific decisions. They are social decisions, and may therefore be over-influenced by hype. Research funding must compete with other funding needs, and research hype is up against the hype of military spending, spiralling oil costs and the latest celebrity exploits. Waiting for medical benefits to be proven before funding basic research would mean that basic research would never be funded.

Medicine is founded on germ theory, physiology, hygiene and antibiotics. Yet there are profound gaps in our understanding of those interactions. Health conditions common in the developed world (obesity, diabetes, allergies, asthma, heart, vascular and inflammatory diseases) remain rare to non-existent in the rural undeveloped world, despite the lack of

potential treatment. The ‘hygiene hypothesis’ tries to explain these differences, but so far no protective agent has been found. If the microbiome project can eventually deliver such an agent, or agents, and stop even a fraction of the health decline associated with economic development, the effects will still be profound.

The analogy of the microbiome with rainforest biodiversity is apt. We should exhaustively sample and understand the diversity of both in the wild before the ‘clear cutting’ of modern agriculture and hygiene practices irreversibly destroys them both.

David R. Whitlock Nitroceutic LLC, Watertown, Massachusetts 02472, USA

e-mail: daedalus4u@yahoo.com

Competing financial interests: D. R. W. has patented the use of autotrophic ammonia-oxidizing bacteria as a topical treatment.

Human microbiome: take-home lesson on growth promoters?

SIR — Your News Feature ‘The inside story’ (*Nature* **453**, 578–580; 2008) highlights the potential importance of the gut microbiome in understanding health. As you discuss, human profiles differ in obese and lean individuals and are being scrutinized for their possible influence on weight.

A factor you don’t mention is the routine administration over some 50 years of antibiotics and probiotics as growth promoters to farm animals. Use of antibiotics for this purpose was banned in 2005 in the European Union, but it continues in many places (including the United States).

Compounds with antibiotic activity and bacteria with probiotic activity have been widely tested as growth promoters (see, for example, M. Khan *et al. Brit. Poultry Sci.* **48**, 732–735; 2007). The probiotics used in agricultural industries are mainly *Lactobacillus*

"Science fiction has become darker, and I think I am one of the people who changed it." Brian Aldiss Q&A, page 698

and *Bifidobacterium* species.

Humans, particularly children, have been taking these same probiotics for many years, especially in fermented dairy products. They also frequently take broad-spectrum antibiotics.

Given the current obesity pandemic among humans and the impact of antimicrobials on weight gain in animals such as poultry and pigs, there may be a case for evaluating the effects of routinely adding bacteria to our food and of long-term consumption of antibiotics.

Didier Raoult Unité des Rickettsies, CNRS UMR 6236, IRD 198, Faculté de Médecine 27, 13385 Marseille, France
e-mail: didier.raoult@gmail.com

Some alphabets easily beat Russian letter count

SIR — In his Essay about the Phaistos Disc ('A century of puzzling' *Nature* **453**, 990–991; 2008), Andrew Robinson notes that the largest known alphabet is Russian, with 36 letters. In fact, the Russian alphabet has had 33 letters since 1918; before that, it officially had 35 (37 in reality).

But even this number falls short by comparison with the alphabets of the many consonant-rich languages of the northern Caucasus. These commonly have more than 40 letters (for example, there are 45 in Lezghin, 49 in Chechen, 51 in Avar and 62 in one of the dialects of Abkhaz).

The outright winner is the Archi alphabet, developed in 2006. This is another language from the Caucasus and has 97 letters — although many of these are groups of two, three or even four or five characters, rather than independent signs. The highest number of independent signs, at 41, is probably to be found in Abkhaz.

Mikhail S. Gelfand Institute for Information Transmission Problems, B. Karetny pereulok 19, Moscow 127994, Russia
e-mail: gelfand@iitp.ru

Medical Research Council values basic research

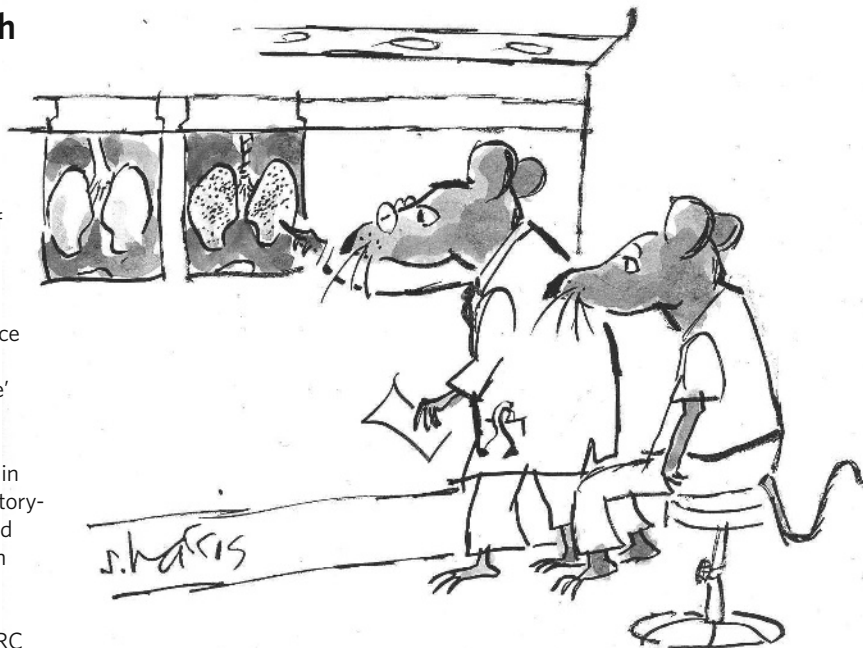
SIR — The Medical Research Council (MRC) recognizes the concerns of some sections of the UK biomedical community, as highlighted by Stephen Moss in his Correspondence 'Translational research: don't neglect basic science' (*Nature* **454**, 274; 2008). The council has been clear that sustained investment in basic research — in laboratory-based as well as clinical and population settings — is an essential foundation for translational research.

The 30% increase in MRC funding over the current spending-review period includes a specific allocation for translational research. However, it also features increased funds for more basic studies. The only significant change for basic researchers is that it will be much easier for them to contribute to translational research and to work related to public health if they wish to do so.

The MRC has always recognized that giving talented investigators scope to pursue their ideas is one of the best ways to advance medical science. It will shortly be announcing the reintroduction of five-year-programme grants to improve support for longer-term research and risk-taking.

The council will continue to support a vibrant and well-resourced science base, acknowledging that investigator-led research, championed within the MRC and throughout the research community, is fundamental to what the MRC does.

Leszek Borysiewicz Medical Research Council, 20 Park Crescent, London W1B 1AL, UK
e-mail: leszek.borysiewicz@headoffice.mrc.ac.uk



Why does work on same mouse models give different results?

SIR — Your News Q&A article 'Lab disinfectant harms mouse fertility' (*Nature* **453**, 964; 2008) must have set tongues wagging in coffee rooms throughout academia. We experienced a similar catastrophe, which took two years and exhaustive detective work to resolve, simply because we moved laboratories.

Our research relies on a widely used model of allergic lung inflammation, in which mice are exposed to a model allergen. While we were working at the University of Cambridge, the model was always reproducible. But when we returned to Ireland to continue this research, we found the lung physiology of the control mice was inexplicably abnormal for the first year.

After months of revising protocols, testing reagents and pathogen screening, we noticed that all the mice developed spontaneous, non-specific lung inflammation within four weeks of arrival, indicating that an environmental

factor was probably to blame.

To cut a long story short, it turned out that mouse chow sterilized by steam autoclaving caused the release of fine particulates, and these were inhaled by mice in their individually ventilated cages. Mice fed instead with irradiated chow had normal lungs.

Unless peer-reviewed and published, such discoveries become anecdotal. So why did we not publish these findings? After a two-year hiatus doing experiments outside the area of our core research, the pressure was on to focus once more on our research aims. Additionally, the growing emphasis on commercially exploitable research, rather than on basic science, means that funding bodies are not impressed by publications on empirical investigations.

This raises a broader, unspoken question in the field of mouse immunology. Why are there differences between data generated by different groups working on apparently the same mouse model or strain?

Padraic G. Fallon Institute of Molecular Medicine, St James's Hospital, Trinity College Dublin, Dublin 8, Ireland
e-mail: pfallon@tcd.ie

"Science fiction has become darker, and I think I am one of the people who changed it." Brian Aldiss Q&A, page 698

and *Bifidobacterium* species.

Humans, particularly children, have been taking these same probiotics for many years, especially in fermented dairy products. They also frequently take broad-spectrum antibiotics.

Given the current obesity pandemic among humans and the impact of antimicrobials on weight gain in animals such as poultry and pigs, there may be a case for evaluating the effects of routinely adding bacteria to our food and of long-term consumption of antibiotics.

Didier Raoult Unité des Rickettsies, CNRS UMR 6236, IRD 198, Faculté de Médecine 27, 13385 Marseille, France
e-mail: didier.raoult@gmail.com

Some alphabets easily beat Russian letter count

SIR — In his Essay about the Phaistos Disc ('A century of puzzling' *Nature* **453**, 990–991; 2008), Andrew Robinson notes that the largest known alphabet is Russian, with 36 letters. In fact, the Russian alphabet has had 33 letters since 1918; before that, it officially had 35 (37 in reality).

But even this number falls short by comparison with the alphabets of the many consonant-rich languages of the northern Caucasus. These commonly have more than 40 letters (for example, there are 45 in Lezghin, 49 in Chechen, 51 in Avar and 62 in one of the dialects of Abkhaz).

The outright winner is the Archi alphabet, developed in 2006. This is another language from the Caucasus and has 97 letters — although many of these are groups of two, three or even four or five characters, rather than independent signs. The highest number of independent signs, at 41, is probably to be found in Abkhaz.

Mikhail S. Gelfand Institute for Information Transmission Problems, B. Karetny pereulok 19, Moscow 127994, Russia
e-mail: gelfand@iitp.ru

Medical Research Council values basic research

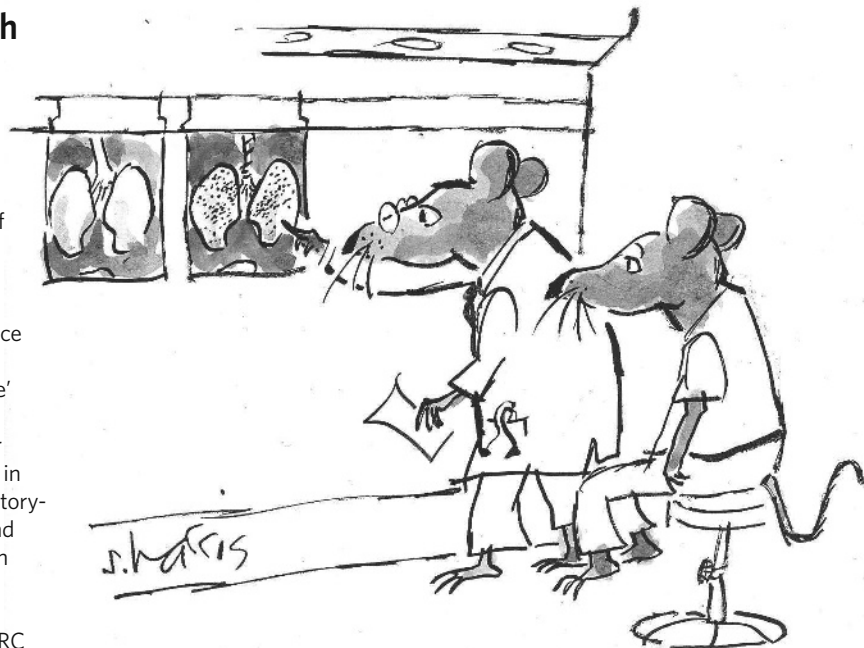
SIR — The Medical Research Council (MRC) recognizes the concerns of some sections of the UK biomedical community, as highlighted by Stephen Moss in his Correspondence 'Translational research: don't neglect basic science' (*Nature* **454**, 274; 2008). The council has been clear that sustained investment in basic research — in laboratory-based as well as clinical and population settings — is an essential foundation for translational research.

The 30% increase in MRC funding over the current spending-review period includes a specific allocation for translational research. However, it also features increased funds for more basic studies. The only significant change for basic researchers is that it will be much easier for them to contribute to translational research and to work related to public health if they wish to do so.

The MRC has always recognized that giving talented investigators scope to pursue their ideas is one of the best ways to advance medical science. It will shortly be announcing the reintroduction of five-year-programme grants to improve support for longer-term research and risk-taking.

The council will continue to support a vibrant and well-resourced science base, acknowledging that investigator-led research, championed within the MRC and throughout the research community, is fundamental to what the MRC does.

Leszek Borysiewicz Medical Research Council, 20 Park Crescent, London W1B 1AL, UK
e-mail: leszek.borysiewicz@headoffice.mrc.ac.uk



Why does work on same mouse models give different results?

SIR — Your News Q&A article 'Lab disinfectant harms mouse fertility' (*Nature* **453**, 964; 2008) must have set tongues wagging in coffee rooms throughout academia. We experienced a similar catastrophe, which took two years and exhaustive detective work to resolve, simply because we moved laboratories.

Our research relies on a widely used model of allergic lung inflammation, in which mice are exposed to a model allergen. While we were working at the University of Cambridge, the model was always reproducible. But when we returned to Ireland to continue this research, we found the lung physiology of the control mice was inexplicably abnormal for the first year.

After months of revising protocols, testing reagents and pathogen screening, we noticed that all the mice developed spontaneous, non-specific lung inflammation within four weeks of arrival, indicating that an environmental

factor was probably to blame.

To cut a long story short, it turned out that mouse chow sterilized by steam autoclaving caused the release of fine particulates, and these were inhaled by mice in their individually ventilated cages. Mice fed instead with irradiated chow had normal lungs.

Unless peer-reviewed and published, such discoveries become anecdotal. So why did we not publish these findings? After a two-year hiatus doing experiments outside the area of our core research, the pressure was on to focus once more on our research aims. Additionally, the growing emphasis on commercially exploitable research, rather than on basic science, means that funding bodies are not impressed by publications on empirical investigations.

This raises a broader, unspoken question in the field of mouse immunology. Why are there differences between data generated by different groups working on apparently the same mouse model or strain?

Padraic G. Fallon Institute of Molecular Medicine, St James's Hospital, Trinity College Dublin, Dublin 8, Ireland
e-mail: pfallon@tcd.ie

COMMENTARY

The science of doping

The processes used to charge athletes with cheating are often based on flawed statistics and flawed logic, says **Donald A. Berry**.

Recently, the international Court of Arbitration for Sport upheld doping charges against cyclist Floyd Landis, stripping him of his title as winner of the 2006 Tour de France and suspending him from competition for two years. The court agreed with the majority opinion of a divided three-member American Arbitration Association (AAA) panel and essentially placed a stamp of approval on a laboratory test indicating that Landis had taken synthetic testosterone. Although Landis asserts his innocence, his options for recourse have all but dried up.

Already, in the run-up to this year's Olympic Games, vast amounts of time, money and media coverage have been spent on sports doping. Several doping experts have contended that tests aren't sensitive enough and let dozens of cheaters slip through the cracks. And some athletes are facing sanctions. Upon testing positive for clenbuterol, US swimmer Jessica Hardy was held back from the Olympic team and faces a two-year ban from the sport. She is attesting her innocence. China has already banned several athletes, some of them for life, on doping charges. Indeed, many world-class athletes will find their life's accomplishments and ambitions, their integrity and their reputations hinging on urine or blood tests. But when an athlete tests positive, is he or she guilty of doping? Because of what I believe to be inherent flaws in the testing practices of doping laboratories, the answer, quite possibly, is no.

In my opinion, close scrutiny of quantitative evidence used in Landis's case show it to be non-informative. This says nothing about Landis's guilt or innocence. It rather reveals that the evidence and inferential procedures used to judge guilt in such cases don't address the question correctly. The situation in drug-testing labs worldwide must be remedied. Cheaters evade detection, innocents are falsely accused and sport is ultimately suffering.

Prosecutor's fallacy

One factor at play in many cases that involve statistical reasoning, is what's known as the prosecutor's fallacy¹. At its simplest level, it concludes guilt on the basis of an observation that would be extremely rare if the person were innocent. Consider a blood test that perfectly matches a suspect to the perpetrator of a crime. Say, for example, the matching profile occurs in just 1 out of every 1,000 people.



P. DEJONG/AP

Floyd Landis (centre) was disqualified after winning the 2006 Tour de France.

A naive prosecutor might try to convince a jury that the odds of guilt are 999:1, that is, the probability of guilt is 0.999. The correct way to determine odds comes from Bayes rule²⁻⁴ and is equal to 999 times $P/(1-P)$ where P is the 'prior probability' of guilt. Prior probability can be difficult to assess, but could range from very small to very large based on corroborating evidence implicating the suspect. The prosecutor's claim that the odds are 999:1 implies a prior probability of guilt equal to 0.5 (in which case P and $1-P$ cancel). Such a high value of P is possible, but it would require substantial evidence. Suppose there is no evidence against the suspect other than the blood test: he was implicated only because he was from the city where the crime occurred. If the city's population is one million then P is 1/1,000,000 and the odds of his guilt are 1001:1 against, which corresponds to a probability of guilt of less than 0.001.

The prosecutor's fallacy is at play in doping cases. For example, Landis's positive test result seemed to be a rare event, but just how rare? In doping cases the odds are dictated by the relative likelihood of a positive test assuming the subject was doping ('sensitivity') against a positive result assuming no doping (which is one minus 'specificity'). Sensitivity and specificity are crucial measures that must be estimated

with reasonable accuracy before any conclusion of doping can be made, in my opinion.

The studies necessary to obtain good estimates are not easy to do. They require known samples, both positive and negative for doping, tested by blinded technicians who use the same procedures under the same conditions present in actual sporting events. In my view, such studies have not been adequately done, leaving the criterion for calling a test positive unvalidated.

Laboratory practices

Urine samples from cyclists competing in the 2006 Tour de France were analysed at the French national anti-doping laboratory (LNDD) in Châtenay-Malabry. This is one of 34 laboratories accredited by the World Anti-Doping Agency to receive and analyse test samples from athletes. The LNDD flagged Landis's urine sample following race stage 17, which he won, because it showed a high ratio of testosterone to epitestosterone.

Based on the initial screening test, the LNDD conducted gas chromatography with mass spectrometry, and isotope ratio mass spectrometry on androgen metabolites in Landis's sample. Such laboratory tests involve a series of highly sophisticated processes that are used to identify the likelihood of abnormal levels of plant-based androgen metabolites (from dietary or

pharmaceutical sources) in a urine sample. The goal is to differentiate from endogenous androgen metabolites normally found in urine.

Mass spectrometry requires careful sample handling, advanced technician training and precise instrument calibration. The process is unlikely to be error-free. Each of the various steps in handling, labelling and storing an athlete's sample represents opportunity for error.

In arbitration hearings, the AAA threw out the result of the LNDD's initial screening test because of improper procedures. In my opinion, this should have invalidated the more involved follow-up testing regardless of whether or not sensitivity and specificity had been determined. Nevertheless, the AAA ruled the spectrometry results sufficient to uphold charges of doping.

During arbitration and in response to appeals from Landis, the LNDD provided the results of its androgen metabolite tests for 139 'negative' cases, 27 'positive' cases, and Landis's stage 17 results (see Fig. 1). These data were given to me by a member of Landis's defence team. The criteria used to discriminate a positive from a negative result are set by the World Anti-Doping Agency and are applied to these results in Fig. 1b and d. But we have no way of knowing which cases are truly positive and which are negative. It is proper to establish threshold values such as these, but only to define a hypothesis; a positive test criterion requires

further investigation on known samples.

The method used to establish the criterion for discriminating one group from another has not been published, and tests have not been performed to establish sensitivity and specificity. Without further validation in independent experiments, testing is subject to extreme biases. The LNDD lab disagrees with my interpretation. But if conventional doping testing were to be submitted to a regulatory agency such as the US Food and Drug Administration⁵ to qualify as a diagnostic test for a disease, it would be rejected.

The problem with multiples

Landis seemed to have an unusual test result. Because he was among the leaders he provided 8 pairs of urine samples (of the total of approximately 126 sample-pairs in the 2006 Tour de France). So there were 8 opportunities for a true positive — and 8 opportunities for a false positive. If he never doped and assuming a specificity of 95%, the probability of all 8 samples being labelled 'negative' is the eighth power of 0.95, or 0.66. Therefore, Landis's false-positive rate for the race as a whole would be about 34%. Even a very high specificity of 99% would mean a false-positive rate of about 8%. The single-test specificity would have to be increased to much greater than 99% to have an acceptable false-positive rate. But we don't know the single-test specificity because the

appropriate studies have not been performed or published.

More important than the number of samples from one individual is the total number of samples tested. With 126 samples, assuming 99% specificity, the false-positive rate is 72%. So, an apparently unusual test result may not be unusual at all when viewed from the perspective of multiple tests. This is well understood by statisticians, who routinely adjust for multiple testing. I believe that test results much more unusual than the 99th percentile among non-dopers should be required before they can be labelled 'positive'.

Other doping tests are subject to the same weak science as testosterone, including tests for naturally occurring substances, and some that claim to detect the presence of a foreign substance. Detecting a banned foreign substance in an athlete's blood or urine would seem to be clear evidence of guilt. But as with testing for synthetic testosterone, such tests may actually be measuring metabolites of the drug that are naturally occurring at variable levels.

Whether a substance can be measured directly or not, sports doping laboratories must prospectively define and publicize a standard testing procedure, including unambiguous criteria for concluding positivity, and they must validate that procedure in blinded experiments. Moreover, these experiments should address factors such as substance used (banned and not), dose of the substance, methods of delivery, timing of use relative to testing, and heterogeneity of metabolism among individuals.

To various degrees, these same deficiencies exist elsewhere — including in some forensic laboratories. All scientists share responsibility for this. We should get serious about interdisciplinary collaborations, and we should find out how other scientists approach similar problems. Meanwhile, we are duty-bound to tell other scientists when they are on the wrong path. ■

Donald A. Berry is head of the Division of Quantitative Sciences, chair of the Department of Biostatistics and Frank T. McGraw Memorial Chair of Cancer Research, MD Anderson Cancer Center, University of Texas, 1400 Pressler Street, Houston, Texas 77030-1402, USA.
e-mail: dberry@mdanderson.org

1. Buchanan, M. The prosecutor's fallacy. *The New York Times* (16 May 2007).
2. Berry, D. A. *Stat. Sci.* **6**, 175–205 (1991).
3. Berry, D. A. *Statistics: A Bayesian Perspective* (Duxbury Press, California, 1996).
4. Berry, D. A. & Chastain, L. A. *Chance* **17**, 5–8 (2004).
5. <http://www.fda.gov/cdrh/osb/guidance/1620.pdf>

Donald Berry testified for the defence team of 1996 Olympian Mary Decker Slaney before a doping hearing board in 1997. He was paid for his time. See Editorial, page 667.

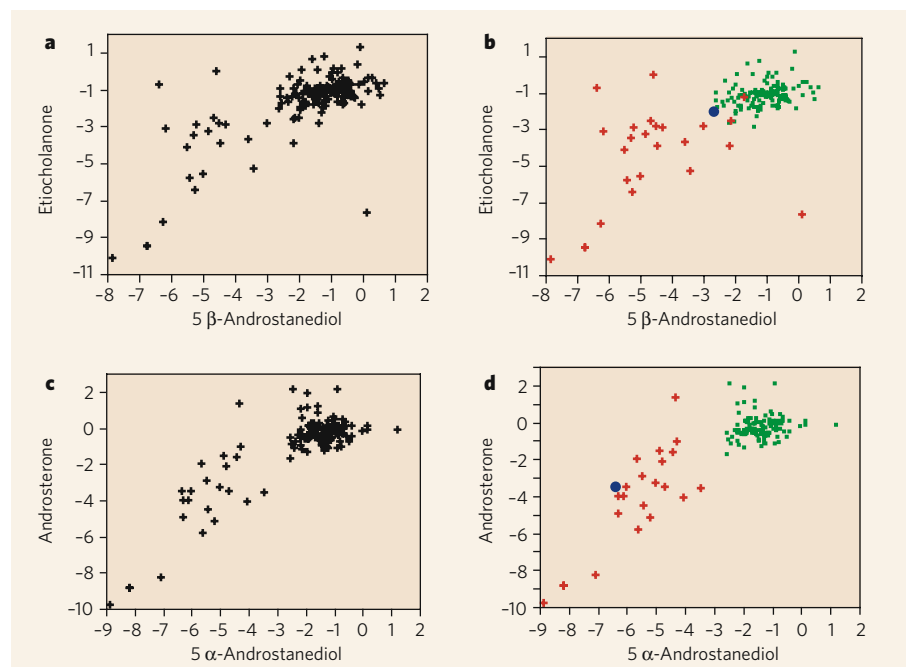


Figure 1 | Metabolite data. Plots show the distribution of 167 samples of the metabolites etiocholanone and 5 β -androstenediol (a, b), and androsterone and 5 α -androstenediol (c, d). Panels b and d show samples the French national anti-doping laboratory (LNDD) designate to be 'positive' (red crosses) or 'negative' (green dots); the values from Landis's second sample from stage 17 is shown as a blue dot. Axes display delta notation, expressing isotopic composition of a sample relative to a reference compound.

BOOKS & ARTS

Powers of observation

A perceptive history documents the many remarkable people who envisioned, built and launched the Hubble Space Telescope, explains **Robert A. Brown**.

The Universe in a Mirror: The Saga of the Hubble Space Telescope and the Visionaries Who Built It

by Robert Zimmerman

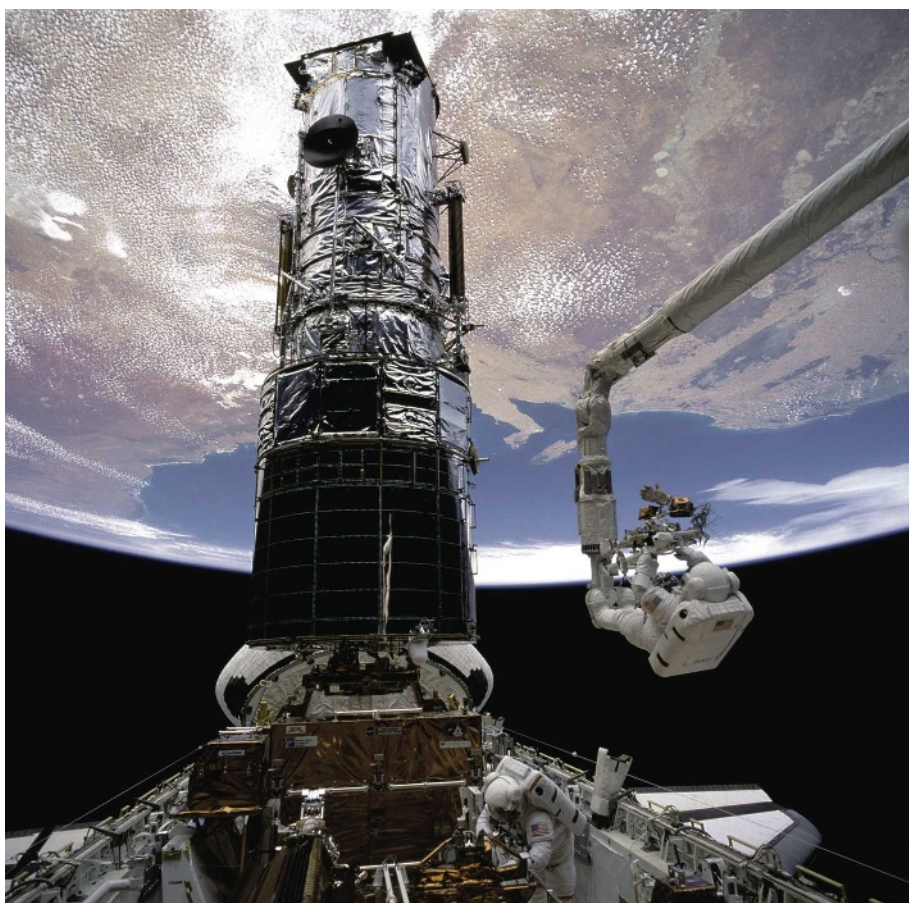
Princeton University Press: 2008.

320 pp. \$29.95, £17.95

Since the winter of 1610, when Galileo discovered the phases of Venus and the moons of Jupiter, telescopes have changed human perspectives. Galileo's telescope proved Copernicus correct — that the Sun, not Earth, lay at the centre of the Solar System, playing no small part in the scientific revolution. Advanced telescopes continue to overturn received wisdom in favour of observational fact. In his history of the Hubble Space Telescope, *The Universe in a Mirror*, science writer and historian Robert Zimmerman states that the telescope “lifted a curtain from our view of the Universe, changing it so profoundly that no human can look at the stars in the same way again”.

The Hubble Space Telescope's cornucopia of breathtaking pictures has shown us the age of the Universe, monster black holes at the hearts of galaxies, vortices roaming the cloud belts of Jupiter, dark matter structuring the Universe, the acceleration of cosmic expansion, and early galaxies merging and growing. The scope and grandeur of Hubble's accomplishments, and the magnificent spectacle of its in-orbit repair and upgrade by heroic astronauts, beg the descriptions that this book superbly answers, about the manner of men and women that did this, how it happened and where it will lead.

Former US president John Quincy Adams may be the only progenitor of the Hubble Space Telescope that Zimmerman does not acknowledge in his telling of the story, which has more dramatis personae than a Russian play. In his first annual message to Congress in 1825, Adams requested public funds for what would have been the first astronomical observatory in the United States. He called it a “lighthouse of the sky”, a concept he pursued throughout his life. Physicist Lyman Spitzer lit his torch from that flame 120 years later, and rightfully takes centre stage in any account of Hubble's origins. Through decades of ups and downs, Spitzer tirelessly promoted the benefits of such a lighthouse — a space telescope with access to the entire spectrum and an image quality limited only by the size of the optic. Even after



NASA

Visionary plan: Hubble was designed to be repaired and upgraded by astronauts while in orbit.

Hubble's launch in 1990, Spitzer was hands-on, helping to devise a comprehensive solution to the tragic problem of spherical aberration in the main mirror, which meant the telescope could not focus adequately. At one point, he conferred daily on the prescription of the coin-size corrective mirrors that astronauts installed in 1993, which fixed the problem perfectly. Zimmerman portrays Spitzer's many dimensions as a technologist, physicist, astronomer, rock climber, visionary and gentleman, revealing a lifetime of outsized opportunities that was tempered with some major disappointments.

The human history behind Hubble is only known because the Smithsonian National Air and Space Museum and the American Institute of Physics arranged for hundreds of players to sit for exhaustive interviews about themselves and their work. Judging from Zimmerman's bibliography, he read them all, and conducted

36 more interviews of his own besides.

The written outcome might have been a pastiche with no convincing verisimilitude to real people, settings and actions. The opposite is true in this book. Zimmerman brings the cast to vibrant life, anchors them in Spitzer's sustaining vision, and sets them to re-enacting their parts in the drama of Hubble struggling towards reality. Each individual, however flawed or gifted, felt they were a part of something big, and fervently hoped for the success of the telescope. They wondered about what Hubble might discover, about our surrounding Universe and our origins in it. For the project managers, engineers, administrators and astronomers alike, Hubble seemed to promise a new relationship with the sky.

If there is a dark character in the book, it is the smoke and mirrors of budgeting and approving high-technology mega-projects.

For years, the official budget for building and operating Hubble for one year was US\$300 million, a number that NASA administrator James Fletcher seemingly “picked out of the air with no connection to actual cost”. Combined with a NASA culture at the time of suppressing problems, this underestimation led to successive budget crunches — but perhaps not, as Zimmerman suggests, to the three washers placed in error on the reflective null corrector of the Hubble test setup at optics company Perkin-Elmer, which ultimately caused the main mirror’s spherical aberration. Not every event has a deep original cause, nor happens because it was not prevented.

We don’t know exactly how the pyramids of Egypt were built; as with Hubble, the planning was probably not linear, and people may well have temporized and played games with the budget and schedule. For better or worse, the planning process works because honest and well-motivated people participate. Fixing

the system, except to better accommodate uncertainties, is not always realistic when unique projects demand the highest performance and operate at the limits of technology.

It is unrealistic to suggest that the scientists, contractors and bureaucrats could have warned Fletcher of the infeasible budget, and either accepted the telescope’s cancellation or forced him to seek more funding. No one knew what Hubble would cost in 1974, and a simple working agreement on a large budget was a form of success. Everyone involved knew where this project was going — they were glad to be part of it and believed that it would be a game-changer.

Hubble has operated for 18 years and counting. Another servicing mission planned for October 2008 will bring it two new instruments and other replacement modules. To the limits of its 2.4-metre aperture, the telescope has revealed the visible Universe to the scientific community and a worldwide public.

What grand vision will motivate the telescope’s successor in the visible-wavelength range? To many, the compelling answer is a Hubble-like telescope some eight times larger, with exponentially greater power to observe fainter objects and in finer detail. Equipped with multiple instruments, including a wide-field camera to detect the wobbling of stars caused by orbiting planets, and a coronagraphic spectrometer to detect atmospheric oxygen and water on faint sources, it could find and characterize potentially habitable planets around hundreds of nearby stars, looking for signs of life. Answering the question ‘Are we alone?’ by means of a telescope is now an obtainable goal, a point that we have reached because of all the people and drama that Zimmerman brings to life in this terrific book. ■

Robert A. Brown is an astronomer at the Space Telescope Science Institute, 3700 San Martin Drive, Baltimore 21218, Maryland, USA.
e-mail: rbrown@stsci.edu

Living Googles?

**Representing Autism:
Culture, Narrative, Fascination**

by Stuart Murray

Liverpool University Press: 2008.

288 pp. £50

Representing Autism examines how we use the word autism and what this reveals about how we think about it. A form of literary criticism or cultural anthropology, this original book fills an important gap. Too often, scientists believe that they have direct, unmediated access to the object of their study. Stuart Murray reminds us that autism is not an unambiguous ‘natural kind’: our scientific taxonomy is also prone to biases.

Autism lies on a spectrum, and comprises two major subgroups: people with classic autism and those with Asperger’s syndrome. These groups share the combination of social-communication difficulties, narrow interests (pejoratively called obsessions) and a love of sameness (known clinically as resistance to change). In classic autism, language development in children is also delayed and they can have additional learning difficulties. Autism and Asperger’s syndrome are genetic in origin, affecting how the brain develops and thus affecting mind and behaviour. The terms ‘autism’ and ‘Asperger’s syndrome’ are medical diagnoses, applied when the defining features interfere with an individual’s ability to function, causing them to suffer.

According to Murray, whether we are making a film or writing a scientific paper about autism, we are superimposing categories on to it. For example, the major charity for families and individuals with autism in Britain, the National Autistic Society, was founded in 1962 as the Society for Psychotic Children. This shift in the name could have affected what we looked for and what we saw. Similar

shifts occurred in the first scientific journal for autism research. Now called the *Journal of Autism and Developmental Disorders*, it began as the *Journal of Autism and Childhood Schizophrenia*. These changes signal how we used to believe autism was just the childhood form of schizophrenia, and how we used to think this condition only affected children.

We now know that autism and schizophrenia are distinct. For example, schizophrenia typically causes delusions, hallucinations and ‘thought disorder’ whereas autism does not. And schizophrenia may entail a rather loose



Dustin Hoffman (seated, centre left) played a card-counting autistic savant in the 1988 film *Rain Man*.

- * Editorials
 - * Research Highlights
 - * Journal Club
 - * News
 - * Column
 - * News Features
 - * Correspondence
 - * Commentary

- * Books and Arts
- * News and Views
- * Feature
- * Articles
- * Letters
- * Corrigendum
- * Technology Features
- * Naturejobs
- * Futures



GLORIOUS GARDENS

The chateau grounds at Chaumont-sur-Loire, France, form the backdrop for a contemporary garden festival, running until October. Sustainability, biodiversity, bees and pollination, taste buds and continental drift are among the ideas brought to life in plants by top landscape designers, mirroring the festival's theme of *partage*, or division and sharing. www.chaumont-jardins.com

ANIMAL PASSIONS

Sexuality in the animal kingdom is the subject of an exhibition running until Spring 2009 at New York City's Museum of Sex. The show, advised by Stanford University ecologist Joan Roughgarden and other scientists, suggests that sex has various roles beyond reproduction and mating. Challenging the simple evolutionary theory of sex selection, Roughgarden aims to change our view of sexual behaviours to a more nuanced one, including orientation and cognition. www.museumofsex.com

HEAVENLY HOUSES

The buildings of influential twentieth-century architect John Lautner are showcased in *Between Earth and Heaven* at the Hammer Museum, Los Angeles, until 12 October. Having mainly worked in California, Lautner famously designed modern homes that are open to the environment, use innovative materials and take on futuristic shapes. www.hammer.ucla.edu



J. WHITE

or casual use of evidence to form beliefs, whereas Asperger's syndrome may be marked by a desire for very tight, unambiguous evidence as a basis for forming beliefs. We now also know that autism is lifelong, and the old view that this was just a condition of childhood has meant there have been few, if any, studies of autism-spectrum conditions in adulthood.

One narrow slice of the autism spectrum disproportionately dominates public perceptions of the conditions. In the 1988 movie *Rain Man*, Dustin Hoffman's character Raymond Babbitt could impressively recall all airline crashes by date of incident, or could tell a waitress her phone number just from seeing her name badge, having memorized the local phone directory up to the letter G. Kim Peek, the real man on whom Babbitt is based, is even more impressive. He can recall every word of every one of the thousands of books he has ever read, and can read two facing pages of a book simultaneously. Interestingly, he also completely lacks a corpus callosum, the connective tissue between the two hemispheres of the brain, which may be related to his remarkable skills. Peek was described as a 'living Google' in a British television documentary about him in February 2006. Another documentary in 2007, entitled *Brain Man*, featured Daniel Tammet, a British man with Asperger's syndrome and synaesthesia, who memorized pi to 22,514 decimal places.

TV documentaries and box-office successes such as *Rain Man* have educated the public only about the savant subtype of autism. This same bias appears in best-selling novels too. The central character Christopher in *The Curious Incident of the Dog in the Night-Time* by Mark Haddon (Jonathan Cape, 2003) was a boy assumed to have Asperger's syndrome who, despite his limited social understanding, took advanced mathematics exams at the age of 13.

In such books and films, it is fiction that ends up educating the largest audience about autism.

Murray discusses why art is attracted to autism. We are fascinated, he posits, because autism seems to strike at the neural systems that define us as a species: the ability to pretend and deceive flexibly, to communicate through hints and innuendo, and to respond with empathy. At the same time, autism can, in some cases, facilitate other systems that also define us as a species, such as the ability to do mathematical calculations. Murray suggests the range of representations of autism in fiction include Spock from *Star Trek* and Sherlock Holmes.

Slanted views about autism can even be found in the research community. On the website of Autism Speaks, the major charity funding autism research in the United States, are the words "This disease has taken our children away. It's time to get them back." This is as clear a statement as one can find of autism as a disease, a view that many but not all autism scientists would endorse. Contrast this with Amanda Baggs's online video *In My Language* (<http://tinyurl.com/2pczl2>), which she launched as a statement about her civil rights as a person with autism, to be recognized and understood as different but not diseased. Another website, Autistics.org, proudly proclaims that people with autism are simply differently wired, and names one of their online groups the Autistic Liberation Front. These statements, along with *Representing Autism*, serve as a valuable reminder that we need to challenge how we conceptualize such medical conditions. ■

Simon Baron-Cohen is director of the Autism Research Centre, University of Cambridge, Cambridge CB2 8AH, UK. His latest book is *Autism and Asperger Syndrome: The Facts*. e-mail: sb205@hermes.cam.ac.uk

Visions of our far future

Year Million: Science at the Far Edge of Knowledge

Edited by Damien Broderick

Atlas: 2008. 330 pp. \$40.00, £20.35

The Victorians' discovery of deep geological time was unsettling. Suddenly, human history was an afterthought, a link between an unthinkable long past and a newly imaginable future. H. G. Wells and his heirs, from Olaf Stapledon and J. D. Bernal to Freeman Dyson, tried to sketch what might come in the millennia after us. Today, most futurists

are preoccupied with the problems of the next ten decades, although a few bright-eyed fabulists still scan the far horizon. Australian author and critic Damien Broderick brings them together in this mind-expanding volume of essays.

A million years, or 40,000 generations, is a long time — but not that long on a cosmic timescale. As Broderick puts it, the number is an emblem of a remote future. Most of the book's contributors pitch their remarks around this target, even though some believe the next few hundred years will see the advent



GLORIOUS GARDENS

The chateau grounds at Chaumont-sur-Loire, France, form the backdrop for a contemporary garden festival, running until October. Sustainability, biodiversity, bees and pollination, taste buds and continental drift are among the ideas brought to life in plants by top landscape designers, mirroring the festival's theme of *partage*, or division and sharing. www.chaumont-jardins.com

ANIMAL PASSIONS

Sexuality in the animal kingdom is the subject of an exhibition running until Spring 2009 at New York City's Museum of Sex. The show, advised by Stanford University ecologist Joan Roughgarden and other scientists, suggests that sex has various roles beyond reproduction and mating. Challenging the simple evolutionary theory of sex selection, Roughgarden aims to change our view of sexual behaviours to a more nuanced one, including orientation and cognition. www.museumofsex.com

HEAVENLY HOUSES

The buildings of influential twentieth-century architect John Lautner are showcased in *Between Earth and Heaven* at the Hammer Museum, Los Angeles, until 12 October. Having mainly worked in California, Lautner famously designed modern homes that are open to the environment, use innovative materials and take on futuristic shapes. www.hammer.ucla.edu



J. WHITE

or casual use of evidence to form beliefs, whereas Asperger's syndrome may be marked by a desire for very tight, unambiguous evidence as a basis for forming beliefs. We now also know that autism is lifelong, and the old view that this was just a condition of childhood has meant there have been few, if any, studies of autism-spectrum conditions in adulthood.

One narrow slice of the autism spectrum disproportionately dominates public perceptions of the conditions. In the 1988 movie *Rain Man*, Dustin Hoffman's character Raymond Babbitt could impressively recall all airline crashes by date of incident, or could tell a waitress her phone number just from seeing her name badge, having memorized the local phone directory up to the letter G. Kim Peek, the real man on whom Babbitt is based, is even more impressive. He can recall every word of every one of the thousands of books he has ever read, and can read two facing pages of a book simultaneously. Interestingly, he also completely lacks a corpus callosum, the connective tissue between the two hemispheres of the brain, which may be related to his remarkable skills. Peek was described as a 'living Google' in a British television documentary about him in February 2006. Another documentary in 2007, entitled *Brain Man*, featured Daniel Tammet, a British man with Asperger's syndrome and synaesthesia, who memorized pi to 22,514 decimal places.

TV documentaries and box-office successes such as *Rain Man* have educated the public only about the savant subtype of autism. This same bias appears in best-selling novels too. The central character Christopher in *The Curious Incident of the Dog in the Night-Time* by Mark Haddon (Jonathan Cape, 2003) was a boy assumed to have Asperger's syndrome who, despite his limited social understanding, took advanced mathematics exams at the age of 13.

In such books and films, it is fiction that ends up educating the largest audience about autism.

Murray discusses why art is attracted to autism. We are fascinated, he posits, because autism seems to strike at the neural systems that define us as a species: the ability to pretend and deceive flexibly, to communicate through hints and innuendo, and to respond with empathy. At the same time, autism can, in some cases, facilitate other systems that also define us as a species, such as the ability to do mathematical calculations. Murray suggests the range of representations of autism in fiction include Spock from *Star Trek* and Sherlock Holmes.

Slanted views about autism can even be found in the research community. On the website of Autism Speaks, the major charity funding autism research in the United States, are the words "This disease has taken our children away. It's time to get them back." This is as clear a statement as one can find of autism as a disease, a view that many but not all autism scientists would endorse. Contrast this with Amanda Baggs's online video *In My Language* (<http://tinyurl.com/2pczl2>), which she launched as a statement about her civil rights as a person with autism, to be recognized and understood as different but not diseased. Another website, Autistics.org, proudly proclaims that people with autism are simply differently wired, and names one of their online groups the Autistic Liberation Front. These statements, along with *Representing Autism*, serve as a valuable reminder that we need to challenge how we conceptualize such medical conditions. ■

Simon Baron-Cohen is director of the Autism Research Centre, University of Cambridge, Cambridge CB2 8AH, UK. His latest book is *Autism and Asperger Syndrome: The Facts*. e-mail: sb205@hermes.cam.ac.uk

Visions of our far future

Year Million: Science at the Far Edge of Knowledge

Edited by Damien Broderick

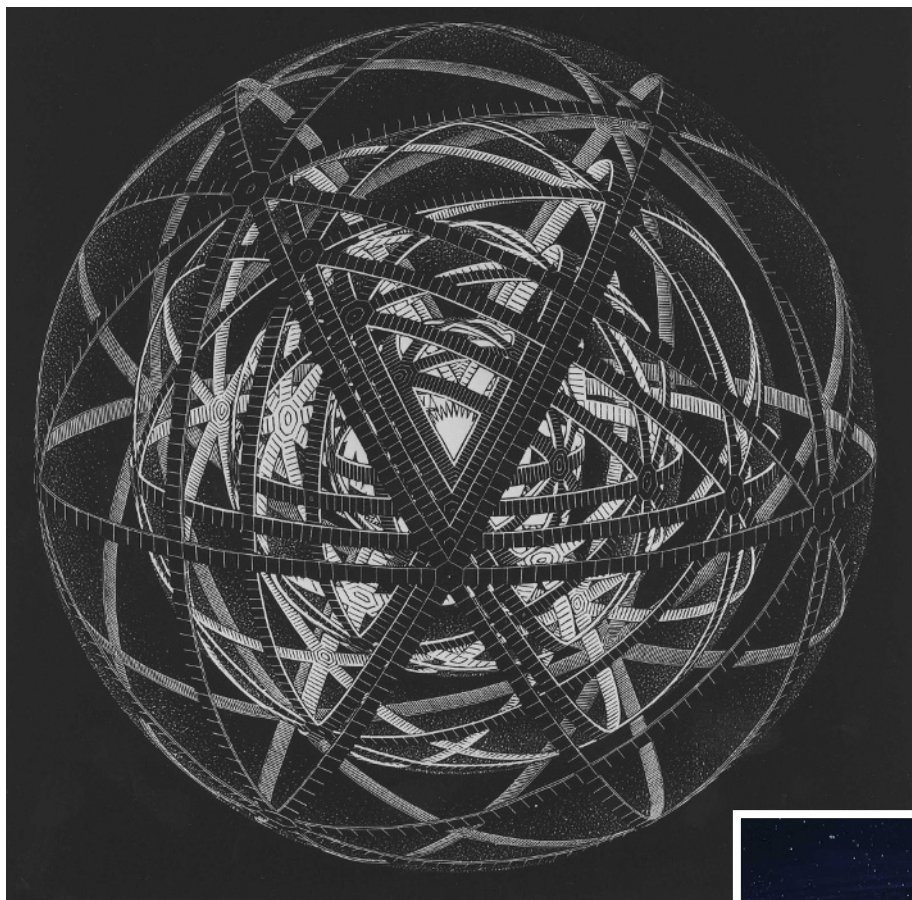
Atlas: 2008. 330 pp. \$40.00, £20.35

The Victorians' discovery of deep geological time was unsettling. Suddenly, human history was an afterthought, a link between an unthinkable long past and a newly imaginable future. H. G. Wells and his heirs, from Olaf Stapledon and J. D. Bernal to Freeman Dyson, tried to sketch what might come in the millennia after us. Today, most futurists

are preoccupied with the problems of the next ten decades, although a few bright-eyed fabulists still scan the far horizon. Australian author and critic Damien Broderick brings them together in this mind-expanding volume of essays.

A million years, or 40,000 generations, is a long time — but not that long on a cosmic timescale. As Broderick puts it, the number is an emblem of a remote future. Most of the book's contributors pitch their remarks around this target, even though some believe the next few hundred years will see the advent

© 2008 THE M. C. ESCHER COMPANY, HOLLAND. ALL RIGHTS RESERVED.



Mirroring M. C. Escher's geometry (above), M-brains could trap a star's energy to power the colonization imagined in H. G. Wells's *First Men In the Moon* (right).

COLUMBIA/KOBAL COLLECTION

of a technological singularity — intelligences of our own devising that go on to develop far beyond ours. They are obliged to write around the thoughts that such intelligences might have, but let us understand that such supercomputer brains will generate these thoughts very, very fast. Only one author, Dougal Dixon, thinks it probable that humans will be extinct within a million years.

The rest of the *Year Million* authors are keener to explore posthuman, extraterrestrial futures that are far removed from the cosy colonialism of *Star Trek*, with its warp drives and humanoid aliens aboard starships. A common feature, in essays from Steven Harris, Wil McCarthy, Robert Bradbury and Robin Hanson, involves humans launching an expanding wavefront of intelligence that engulfs the Galaxy, if not the Universe. Aliens are tactfully assumed to be absent so this great expansion can be reassuringly free of atrocity. This is just as well, because their planets will be converted into generic nano-engineered stuff known as 'computronium'. Thus transformed, this smart hypothetical material

will be assembled into Matrioshka brains, or M-brains. These are the current incarnation of Dyson shells, first described in detail by Freeman Dyson in 1959, which are built round a star to collect all its radiant energy. The M-brains, like nested Russian dolls, would be a series of concentric Dyson shells. Each shell would be filled with intelligence, using waste energy from the one beneath.

Earth escapes the attentions of these 'matter miners' for sentimental reasons. But nowhere else is safe. The stripped-down Darwinism shaping these essays means that more is better. The M-brains send out not starships but seeds, which are efficient and encode the ability to bootstrap new nano-manufacturing capacity on any world that happens to be suitable. Minds can then take up residence, either

downloaded from the seed or transmitted through the, ahem, 'Universenet'.

All this is presented as highly plausible, but any reader will have doubts at some point, depending on their view of how far current technological trajectories can be sustained. As Broderick promises at the outset, these far-flung fantasies are a lot of fun. In this context, Lisa Kaltenegger's detailed look at the near-term chances of pinpointing Earth-like extrasolar planets, one of the most interesting scientific prospects of our times, comes across as slightly prosaic.

In general, the science of the subtitle is a badge of allegiance. Of course we cannot know what future science will reveal. But these are futures that science and technology might bring us. Religion does not get a look in, and deep motives or ultimate meaning are not really up for discussion either. Expansion rules the day. The authors, mostly American, have a hands-on, can-do attitude. Show them a body, a planet, a star system, maybe even a universe, and they will plan how to re-engineer it.

These are not futures of quiet contemplation. Rudy Rucker's engaging vision of an era when evolved intelligence will tap into the quantum computation of the cosmos to

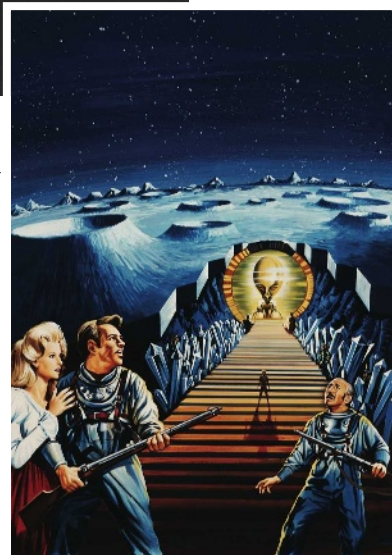
let us "commune with the souls of stones" is a startling exception.

As the horizon shifts far beyond a million years, the cosmological outlook is more familiar. Now that the Big Crunch, with its hint of rebirth as the Universe collapses into a black hole, seems to have been ruled out, the prospect is of an attenuated, etiolated cosmos, an endless vista of cold and dark.

Various get-out clauses invoking the multiverse are canvassed, in which we merely inhabit one

inflationary bubble, or one of many alternative dimensions. Those aside, the final outcome recalls another Victorian preoccupation, the ultimate heat death of the Universe. In *Year Million*, these twenty-first-century visionaries foresee virtually infinite possibilities along the way. ■

Jon Turney is a science writer based in Bristol, UK. His next book, *The Rough Guide to the Future*, will be published in 2009.
e-mail: jonturney@googlemail.com



Q&A: Turning up the heat on sci-fi

The novel *Hothouse* is **Brian Aldiss's** extraordinary 1962 vision of rampant global warming, set on a future stationary Earth with one side permanently baked by the Sun. As the book is republished as a Penguin Modern Classic, the 83-year-old author shares his thoughts on science fiction and life with *Nature*.

What sparked your vision of vigorous plants taking over Earth in *Hothouse*?

My memories of the overwhelming exuberance of vegetation I saw in south and east Asia, where I was in action in the Second World War. The Sumatran jungle captured my imagination. I also remember a vast banyan tree in the Calcutta Botanical Gardens, said to be the biggest tree in the world. The power and energy of these plants was amazing.

Are you concerned about climate change?

The situation is awful. To think we may lose the polar bears. And the idea that bees are in peril. If the bees go, there goes fertilization. We follow.

Climate change is very difficult to deal with because there's no carrot at the end of it. The temptation for present generations to hope future generations will sort it out with a technological fix is very strong.

When I wrote *Hothouse*, apart from the threat of nuclear war, the burning issue was population explosion. I'm surprised to see so little mention of this today. Perhaps it is because a sense of responsibility is lacking in the West: there may be fewer of us, but we are greedier and contribute much more per head to global heating than those in poorer regions.

How would you counter critics who argue that the book is scientific nonsense, for instance, attaching Earth to the Moon by giant cobwebs?

It is science as fantasy. My premise for Earth and the Moon being stationary is that there is a 'trojan' position in which three heavenly bodies — in this case the Sun, Earth and the Moon — get locked into a position from which they cannot escape. I do not claim that Earth and the Moon have been fixed in position by spider webs between them; that's a widely held misunderstanding.

Neil Gaiman's introduction says that the book is less optimistic than US science fiction of the same period. Do you agree?

British science fiction is cautionary. You could argue we derived that from



Sumatran plants such as *Rafflesia arnoldii* inspired Brian Aldiss (below left) to write his futuristic sci-fi novel *Hothouse*.



H. G. Wells. His 1898 novel *War of the Worlds* is a homily against human arrogance. But the typical US mentality is that 'we can sort this out'. Garrett P. Serviss wrote a sequel to Wells's book called *Edison's Conquest of Mars*, in which Americans go to Mars and clobber

the martians. To some extent this thinking has changed. Science fiction has become darker, and I think I am one of the people who changed it.

What inspired you to write?

As a child I had to live within my mother's awful delusion. I was 50 years old before I found out that my sister, who my mother said had died at the age of five, had in fact never lived — she was stillborn. That's the sort of thing that provokes one to invent many paradoxical worlds. I have always enjoyed inventing planets that contrast with Earth.

As a young man, coming back from the East after the end of the Second World War, I sat up on the deck of the troopship watching the bow waves and read Homer's *Odyssey*. It brought me back to worlds of

creativity. I particularly enjoy such stories of the Bronze Age just emerging into the Iron Age, as in my 2004 novel *Jocasta*.

What influences your writing now?

My day-to-day life is very happy. Each morning I get up and am almost immediately excited by some inspiring idea. I'm reading six books at once, which is absurd! There was one more I wanted to read — *Our Final Century* by Martin Rees — but it's so pessimistic I couldn't face it.

I am attracted to what is most remote. In 2015, the NASA probe New Horizons will send back data from Pluto. We'll discover so much about what it is really like. I hope I live to see or hear about it all, even from a hospital bed. The probe will then go on to the Oort Cloud near the edge of the Solar System. Imagine!

Interview by **Caspar Henderson**, a writer based in Oxford, UK. He is writing *A Book of Barely Imagined Beings*.

e-mail: caspar81@gmail.com

Hothouse

by Brian Aldiss

Penguin: 2008. 288 pp. Can\$19, £8.99

Hidden treasures: Padua's anatomy theatre

Alison Abbott finds that human dissections during the Renaissance were rather respectable after all.

When passing through Padua in 1786, Johann Wolfgang von Goethe was surprised to find the Italian university's famed anatomical theatre small and dark. "The audience are squashed on top of each other in a tall pointed funnel," he wrote in his book *Italian Journey*. "They look down the steep sides to a narrow floor where there is a table on which no daylight falls, so the lecturer has to demonstrate by the lamplight." Every winter between 1595 and 1872, this curious wooden structure was used by anatomy professors to dissect the corpses of prisoners condemned to the gallows.

Temporary and makeshift anatomy theatres had already been put to use for more than a century as Renaissance artists and scientists — Leonardo da Vinci himself partook — tried to understand for themselves how the body was put together. But that in Padua is Europe's oldest surviving permanent anatomical theatre, and rises through two storeys of the fourteenth-century Bo Palace, the central building of the University of Padua, some 35 kilometres west of Venice. Around 200 students, along with local dignitaries who were granted the best seats, would squeeze into the theatre's elegantly carved oval tiers, each barely 40 centimetres wide. That the structure could bear such concentrated weight is testament to the local skills in wood-working, honed by the demands of shipbuilding. The Republic of Venice made its immense wealth from marine trading.

Contrary to myth, there is no tunnel below the theatre through which bodies were hidden from Vatican spies. It is true that some popes interpreted a papal bull from 1300 as forbidding dissection, although the charter's intention was to stop the practice of cutting up Crusaders' bodies to send them home in a convenient way. But human dissection was legal in the Republic of Venice, far from Rome. The church insisted only that everything was properly administrated and that corpse remains were given a church blessing afterwards.

However, administration could be slow, and medical students demanding. In 1556 Gabriele Falloppio, who described the Fallopian tubes that now bear his name, lobbied magistrates to

provide more corpses in case Padua lost students to competing universities in Bologna or Ferrara. Even when the permanent theatre opened at the end of the sixteenth century, official dissections never kept pace with academic need, so unofficial ones were held for smaller audiences in professors' homes or in pharmacies. These dissections were also legal, but students may have occasionally stolen corpses to meet demand.

Musicians were brought in to the theatre to entertain during breaks. After all, a full dissection could take a good two weeks — also the reason why dissections were only tolerable in the cold months. Out of Christian respect, they stopped at the beginning of Lent, 40 days before Easter, whenever it happened to fall.

Dissections were cutting-edge science in Renaissance Italy, and Padua was a major international centre for all disciplines. Its ancient university drew in students and professors from around Europe. These included Poland's Nicolaus Copernicus, who worked out that Earth goes around the Sun rather than the other way around, and the Englishman William Harvey, who deduced how

blood circulates in the body. Galileo acquired his first telescope and pointed it to the heavens in Padua, describing his time there as "the eighteen best years of my life". And then, of course, there was Andreas Vesalius from Brussels. Working with an illustrator who was a pupil of the Italian Renaissance painter Titian, Vesalius produced his masterpiece *De Humani Corporis Fabrica* (On the Fabric of the Human Body) in 1543. It was as radical as Copernicus's *De Revolutionibus* (On the Revolutions), published in the same year, and it excited further interest in dissection, encouraging the construction of many permanent anatomical theatres around Europe.

These days the anatomical theatre serves only as a museum piece, open to the public for guided tours. But a small hall next to it is still used as a meeting room for medical faculty.

The march of centuries is acknowledged here in the sombre display of skulls of noted professors, exhumed during the phrenology frenzy of the 1800s, and in the twentieth-century frescoes complementing those from the sixteenth century. The university has

other historical science collections, including its splendid and unique array of eighteenth-century physics instruments. It also has a charming Renaissance botanical garden, as surprisingly compact as the anatomical theatre — and also the oldest of its kind. Created in 1545, the garden was used to teach students about medicinal plants. Its longest surviving tree is the Goethe palm, planted in 1585. The palm so impressed the great writer-scientist during his Padua sojourn 200 years later that he described it in his 1790 seminal text *Metamorphosis of Plants*. ■

Alison Abbott is Nature's Senior European correspondent.

See <http://tinyurl.com/56w942> and www.musei.unipd.it/en/index.html for further details.



Padua boasts an anatomy theatre and botanical garden, both Europe's oldest.



M. DANESIN

NEWS & VIEWS

EARTH SCIENCE

Structuring the inner core

John Lister

Earth's rocky mantle and solid inner core are separated by the 2,300-kilometre-deep layer of molten iron that constitutes the outer core. Yet the sluggish pattern of mantle convection creates structure in the inner core.

In the past decade or so, seismologists have discovered surprising differences between the elastic properties of the eastern and western hemispheres of Earth's solid inner core^{1,2}. In parallel, studies of Earth's changing magnetic field reveal persistent patterns of flow in the liquid outer core³, where the field is generated by dynamo action. The field itself has persistent features that are not symmetric about Earth's axis of rotation (which is why the north and south magnetic poles differ from the geographic poles).

On page 758 of this issue, Aubert *et al.*⁴ show how these very different observations are related to the pattern of heat flux across the boundary between the core and the mantle. Their calculations provide a satisfyingly coherent explanation based on dynamic thermal connections between Earth's concentric shells. Indirectly, it seems, the history of tectonic-plate motion on Earth's surface may have been frozen into structure at the centre of the Earth.

The story begins with two coupled convecting layers (Fig. 1). The upper layer, the mantle, cools as heat passes through the lithosphere, oceans and atmosphere to space, and gains some heat from the lower, hotter layer, the outer core. Convection in the mantle is very slow owing to its great viscosity. Typical velocities, seen in the motions of the surface plates, are only a few centimetres per year, and the arrangement of downwellings and upwellings — subduction zones and plumes — changes slowly with a typical timescale of a few hundred million years.

By contrast, the liquid outer core convects much more rapidly as it loses heat to the mantle. Its viscosity is debated — something comparable to that of water is possible — but typical velocities certainly reach a few centimetres per minute. Flow in the electrically conducting outer core rearranges the magnetic field, causing changes in the field at Earth's surface. By analysing historical records of the surface field, one can deduce the size and pattern of flow at the top of the core⁵. It fluctuates on decadal timescales, but some persistent features, such as a cyclonic circulation beneath central Asia,

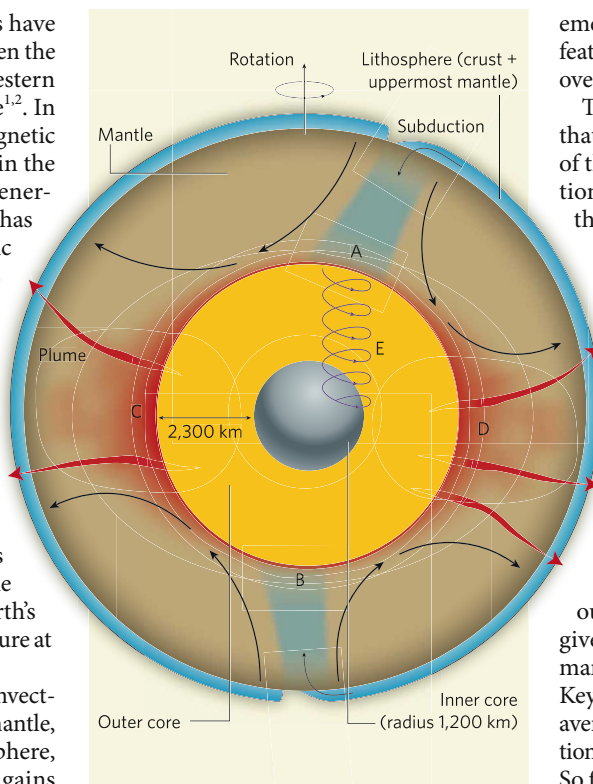


Figure 1 | Coupled convection in the Earth.

Large-scale convection in the mantle is related to lithospheric plate motion and the configuration of subduction zones, where cold, dense material slowly sinks from the surface towards the core-mantle boundary. The heat flux from the outer core is much greater beneath cold downwellings (A, B) than beneath warm upwelling plumes (C, D). The more rapid convection in the liquid outer core is driven by cooling from the mantle and by release of light fluid at the solidifying inner-core boundary. Earth's rotation tends to organize this flow into helical convection columns (E), which allow the pattern of mantle cooling to affect the long-term pattern of inner-core solidification. Present-day subduction occurs predominantly around the Pacific 'rim of fire', giving asymmetry between the eastern and western hemispheres. Aubert and colleagues' simulations⁴ of flow in the outer core combine the effects of inner-core solidification with those of mantle heat-flux variations to provide a unifying explanation for seismological and geomagnetic data. (Modified from Fig. 7 of ref. 9.)

emerge in a 150-year time average. Are these features controlled by spatial variations in the overlying, almost stationary mantle?

To answer this question, Aubert *et al.*⁴ reason that variations in seismic velocity at the base of the mantle probably correspond to variations in temperature, not composition, and thus to variations in the heat flux from the core. Where cooler mantle has recently descended to the core-mantle boundary ('recently', that is, on the 100-million-year mantle timescale), the heat flux will be many times that where mantle material has been warmed by the core for a long time.

Because of the huge difference in viscosities, the mantle sets the heat flux, and the core responds. So Aubert *et al.* apply the pattern of heat flux deduced from mantle seismology⁶ to a numerical model of flow and dynamo action in the outer core. Reassuringly, their calculations give magnetic fields and flows at the core-mantle boundary resembling those of Earth. Key to this successful comparison is the time-averaging used to remove the decadal fluctuations and reveal the longer-term mantle control. So far, so good. But what about those intriguing hemispheric differences in the inner core, 2,300 km below the core-mantle boundary?

The major innovation introduced by Aubert *et al.* is to add the effects of inner-core solidification to those of mantle heat-flux variations. The outer-core liquid is an alloy of iron and lighter elements. On solidification, the iron is preferentially incorporated into the inner core, leaving behind liquid enriched in the lighter elements. This production of light liquid at the inner-core boundary adds a significant compositional contribution to convection. The local rate of solidification is determined by the flux of heat and composition at the inner-core boundary, and thus itself depends on the pattern of convection that it helps to drive.

One might naively expect that the regions of most rapid inner-core solidification should lie roughly beneath the regions of greatest mantle cooling. But that neglects the strong effect of Earth's rotation, which organizes the outer-core convection into helical flows along columnar

cells aligned parallel to the rotation axis and at a tangent to the inner core. Thus, a prediction from the strong cooling above the central Asian cyclone in the calculations is that there will be greater solidification rates in a broad equatorial region beneath Malaysia, exactly where seismologists locate the eastern hemispheric peak in seismic velocity. Moreover, laboratory experiments on solidifying ingots⁷ show that rapid solidification gives a crystal fabric — less oriented, finer grained — that Aubert *et al.*⁴ argue is also consistent with the greater seismic isotropy and attenuation observed in the eastern hemisphere.

Case proved? Perhaps — although I find that it takes an eye of faith to focus on the similarities beneath central Asia and overlook dissimilarities, such as the strong anticyclone and flux patch south of Africa that are evident in the Earth but not in Aubert and colleagues' model. And it is an interesting question whether the variations with solidification rate in small-scale laboratory experiments⁷ can be scaled up and slowed down to the conditions of inner-core growth.

Aubert *et al.*⁴ estimate that, to solidify a 100-km-thick layer with seismically observable east–west differences, the pattern of solidification rates must be maintained for at least 100 million to 300 million years — and this is for a young and rapidly growing inner core. Over this period, the pattern of mantle cooling must be maintained, and the inner core held almost fixed relative to the mantle by gravitational coupling (allowing small torsional oscillations about equilibrium). The timing is tight, but seems feasible.

A fascinating implication of the picture provided by Aubert *et al.*⁴ is that major differences in the configurations of the plates and the pattern of mantle convection in the past could have been imprinted on the inner core at earlier stages of its growth. Perhaps the inner-core structure is something like an onion, with different layers recording different mantle epochs. Could seismologists 'peel the onion' to reveal evidence of the history of both inner-core growth and mantle convection? Newly published images of the inner core

reveal a seismically distinct 'inner inner core' of radius 600 km, outside which the eastern hemisphere, but not the western, is layered⁸. The calculations of Aubert *et al.* offer a new way to interpret such images, and give insight into this most central of problems in Earth's evolution.

John Lister is at the Institute of Theoretical Geophysics, Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Cambridge CB3 0WA, UK. e-mail: lister@damtp.cam.ac.uk

1. Tanaka, S. & Hamaguchi, H. *J. Geophys. Res.* **102**, 2925–2938 (1997).
2. Creager, K. C. *J. Geophys. Res.* **104**, 23127–23139 (1999).
3. Amit, H. & Olson, P. *Phys. Earth Planet. Inter.* **155**, 120–139 (2006).
4. Aubert, J., Amit, H., Hulot, G. & Olson, P. *Nature* **454**, 758–761 (2008).
5. Bloxham, J. & Jackson, A. *Rev. Geophys.* **29**, 97–120 (1991).
6. Masters, G. *et al.* in *Earth's Deep Interior* Vol. 117 (eds Karato, S. *et al.*) 63–87 (Am. Geophys. Union, 2000).
7. Bergman, M. I., Agrawal, S., Carter, M. & Macleod-Silberstein, M. *J. Crystal Growth* **255**, 204–211 (2003).
8. Sun, X. & Song, X. *Earth Planet. Sci. Lett.* **269**, 56–65 (2008).
9. Gonnermann, H. M., Jellinek, A. M., Richards, M. A. & Manga, M. *Earth Planet. Sci. Lett.* **226**, 53–67 (2004).

TUBERCULOSIS

Shrewd survival strategy

Steven A. Porcelli

***Mycobacterium tuberculosis* modulates its virulence to cause persistent but often subclinical infection. This strategy is regulated in part by a feedback loop that controls the secretion of a small subset of bacterial proteins.**

The bacterium that causes tuberculosis is one of the most successful pathogens. Its spread among humans has been so efficient that as much as one-third of the world's population is now believed to be infected¹. In most cases, these infections cause clinically silent disease, which is likely to remain permanently dormant unless the host's immunity is substantially

compromised. The extraordinary stealth and opportunism that *Mycobacterium tuberculosis* exhibits results from the complex and delicately coordinated way in which it interacts with its host — a process controlled in part by a specialized bacterial protein-secretion system called ESX-1. On page 717 of this issue, Raghavan *et al.*² provide insight into how mycobacteria

might maintain ESX-1 activity at just the right level, perhaps allowing these pathogens to avoid excessive virulence so that they can patiently pursue their low-key yet incredibly effective strategy for long-term survival.

Mycobacteria lack the specialized type I–VI secretion systems that have been well characterized in Gram-negative bacteria. Instead, their virulence is mediated in part by ESX-1, which is also found in other Gram-positive bacteria. Two proteins secreted by the ESX-1 system, ESAT-6 and CFP-10, are prominent targets of the immune system in animals, including humans, infected with *M. tuberculosis*³. These proteins are encoded together as part of a coordinated unit of genetic material, the *esxB* operon. Genomic studies⁴ predicted that genes surrounding *esxB* are likely to encode components of a secretory apparatus responsible for the export of ESAT-6 and CFP-10 out of the bacterial cytoplasm. This prediction was soon validated^{5,6}, defining ESX-1 as a prototypical virulence-associated mycobacterial protein-secretion system.

Exactly how ESX-1 contributes to mycobacterial virulence remains a puzzle, although various observations have led to several hypotheses that are not mutually exclusive. For example, ESX-1 mutants grow poorly during the early phase of infection in mice and in macrophages^{6,7} (immune cells that engulf pathogens by phagocytosis). Moreover, in infected dendritic cells (another type of phagocytic immune cell), ESX-1 helps *M. tuberculosis* escape from phagosome vesicles — in which they are captured for destruction — into the cytoplasm, where they can replicate⁸. Because purified ESAT-6 acts as a membrane-disrupting toxin *in vitro*⁹, this protein could directly mediate phagosomal escape. However, ESX-1 also

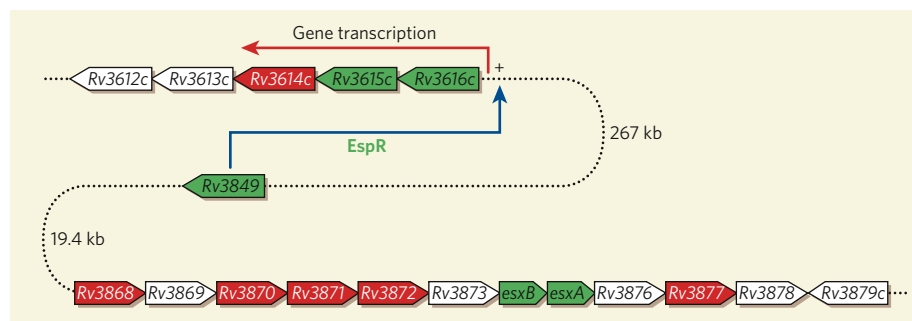


Figure 1 | Genes involved in ESX-1 secretion. At least three separate regions of the *M. tuberculosis* chromosome encode genes that are involved in the ESX-1-secretion system. The *esxB* operon encodes the main secreted products ESAT-6 and CFP-10. Raghavan *et al.*² identify the protein product of the *Rv3849* gene — which is some distance (19.4 kilobases) away from the *esxB* operon — as the EspR protein. This protein is both a secreted substrate of ESX-1 and a positive transcriptional regulator of a third region, the *Rv3616c–Rv3612c* operon, which is also involved in ESX-1-mediated virulence. Genes encoding known secreted substrates of ESX-1 are shown in green, and those encoding structural or functional components of the ESX-1 secretion apparatus are in red. Genes in white are of unknown function, although they are presumed to be somehow involved in the ESX-1 system.

cells aligned parallel to the rotation axis and at a tangent to the inner core. Thus, a prediction from the strong cooling above the central Asian cyclone in the calculations is that there will be greater solidification rates in a broad equatorial region beneath Malaysia, exactly where seismologists locate the eastern hemispheric peak in seismic velocity. Moreover, laboratory experiments on solidifying ingots⁷ show that rapid solidification gives a crystal fabric — less oriented, finer grained — that Aubert *et al.*⁴ argue is also consistent with the greater seismic isotropy and attenuation observed in the eastern hemisphere.

Case proved? Perhaps — although I find that it takes an eye of faith to focus on the similarities beneath central Asia and overlook dissimilarities, such as the strong anticyclone and flux patch south of Africa that are evident in the Earth but not in Aubert and colleagues' model. And it is an interesting question whether the variations with solidification rate in small-scale laboratory experiments⁷ can be scaled up and slowed down to the conditions of inner-core growth.

Aubert *et al.*⁴ estimate that, to solidify a 100-km-thick layer with seismically observable east–west differences, the pattern of solidification rates must be maintained for at least 100 million to 300 million years — and this is for a young and rapidly growing inner core. Over this period, the pattern of mantle cooling must be maintained, and the inner core held almost fixed relative to the mantle by gravitational coupling (allowing small torsional oscillations about equilibrium). The timing is tight, but seems feasible.

A fascinating implication of the picture provided by Aubert *et al.*⁴ is that major differences in the configurations of the plates and the pattern of mantle convection in the past could have been imprinted on the inner core at earlier stages of its growth. Perhaps the inner-core structure is something like an onion, with different layers recording different mantle epochs. Could seismologists 'peel the onion' to reveal evidence of the history of both inner-core growth and mantle convection? Newly published images of the inner core

reveal a seismically distinct 'inner inner core' of radius 600 km, outside which the eastern hemisphere, but not the western, is layered⁸. The calculations of Aubert *et al.* offer a new way to interpret such images, and give insight into this most central of problems in Earth's evolution.

John Lister is at the Institute of Theoretical Geophysics, Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Cambridge CB3 0WA, UK. e-mail: lister@damtp.cam.ac.uk

1. Tanaka, S. & Hamaguchi, H. *J. Geophys. Res.* **102**, 2925–2938 (1997).
2. Creager, K. C. *J. Geophys. Res.* **104**, 23127–23139 (1999).
3. Amit, H. & Olson, P. *Phys. Earth Planet. Inter.* **155**, 120–139 (2006).
4. Aubert, J., Amit, H., Hulot, G. & Olson, P. *Nature* **454**, 758–761 (2008).
5. Bloxham, J. & Jackson, A. *Rev. Geophys.* **29**, 97–120 (1991).
6. Masters, G. *et al.* in *Earth's Deep Interior* Vol. 117 (eds Karato, S. *et al.*) 63–87 (Am. Geophys. Union, 2000).
7. Bergman, M. I., Agrawal, S., Carter, M. & Macleod-Silberstein, M. *J. Crystal Growth* **255**, 204–211 (2003).
8. Sun, X. & Song, X. *Earth Planet. Sci. Lett.* **269**, 56–65 (2008).
9. Gonnermann, H. M., Jellinek, A. M., Richards, M. A. & Manga, M. *Earth Planet. Sci. Lett.* **226**, 53–67 (2004).

TUBERCULOSIS

Shrewd survival strategy

Steven A. Porcelli

***Mycobacterium tuberculosis* modulates its virulence to cause persistent but often subclinical infection. This strategy is regulated in part by a feedback loop that controls the secretion of a small subset of bacterial proteins.**

The bacterium that causes tuberculosis is one of the most successful pathogens. Its spread among humans has been so efficient that as much as one-third of the world's population is now believed to be infected¹. In most cases, these infections cause clinically silent disease, which is likely to remain permanently dormant unless the host's immunity is substantially

compromised. The extraordinary stealth and opportunism that *Mycobacterium tuberculosis* exhibits results from the complex and delicately coordinated way in which it interacts with its host — a process controlled in part by a specialized bacterial protein-secretion system called ESX-1. On page 717 of this issue, Raghavan *et al.*² provide insight into how mycobacteria

might maintain ESX-1 activity at just the right level, perhaps allowing these pathogens to avoid excessive virulence so that they can patiently pursue their low-key yet incredibly effective strategy for long-term survival.

Mycobacteria lack the specialized type I–VI secretion systems that have been well characterized in Gram-negative bacteria. Instead, their virulence is mediated in part by ESX-1, which is also found in other Gram-positive bacteria. Two proteins secreted by the ESX-1 system, ESAT-6 and CFP-10, are prominent targets of the immune system in animals, including humans, infected with *M. tuberculosis*³. These proteins are encoded together as part of a coordinated unit of genetic material, the *esxB* operon. Genomic studies⁴ predicted that genes surrounding *esxB* are likely to encode components of a secretory apparatus responsible for the export of ESAT-6 and CFP-10 out of the bacterial cytoplasm. This prediction was soon validated^{5,6}, defining ESX-1 as a prototypical virulence-associated mycobacterial protein-secretion system.

Exactly how ESX-1 contributes to mycobacterial virulence remains a puzzle, although various observations have led to several hypotheses that are not mutually exclusive. For example, ESX-1 mutants grow poorly during the early phase of infection in mice and in macrophages^{6,7} (immune cells that engulf pathogens by phagocytosis). Moreover, in infected dendritic cells (another type of phagocytic immune cell), ESX-1 helps *M. tuberculosis* escape from phagosome vesicles — in which they are captured for destruction — into the cytoplasm, where they can replicate⁸. Because purified ESAT-6 acts as a membrane-disrupting toxin *in vitro*⁹, this protein could directly mediate phagosomal escape. However, ESX-1 also

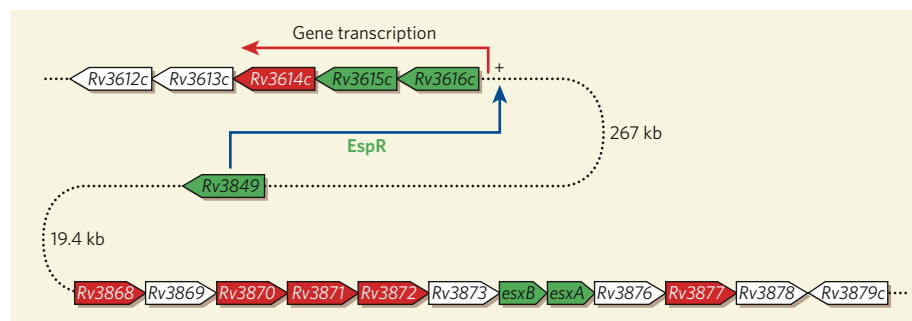


Figure 1 | Genes involved in ESX-1 secretion. At least three separate regions of the *M. tuberculosis* chromosome encode genes that are involved in the ESX-1-secretion system. The *esxB* operon encodes the main secreted products ESAT-6 and CFP-10. Raghavan *et al.*² identify the protein product of the *Rv3849* gene — which is some distance (19.4 kilobases) away from the *esxB* operon — as the EspR protein. This protein is both a secreted substrate of ESX-1 and a positive transcriptional regulator of a third region, the *Rv3616c–Rv3612c* operon, which is also involved in ESX-1-mediated virulence. Genes encoding known secreted substrates of ESX-1 are shown in green, and those encoding structural or functional components of the ESX-1 secretion apparatus are in red. Genes in white are of unknown function, although they are presumed to be somehow involved in the ESX-1 system.

hinders phagosome maturation in macrophages through a process that does not require ESAT-6 or other known substrates secreted by this system¹⁰. Taken together, these data indicate that ESX-1 modulates innate immune responses of the infected host through several mechanisms, which probably involve ESX-1-mediated secretion of ESAT-6 and other unidentified factors.

Raghavan and colleagues² identify a previously unknown component of ESX-1 that is not only a central regulator of it but is also secreted by this system. They find that, in *M. tuberculosis*, the *Rv3849* gene — which is located some distance from the main ESX-1 gene cluster surrounding the *esxB* operon — is required for ESX-1 function. The protein product of *Rv3849*, EspR, is highly similar to a gene transcription factor of the harmless soil bacterium *Bacillus subtilis*. The authors find that EspR is also a DNA-binding transcriptional regulator — *Rv3849* deactivation leads to changes in the transcription of a few operons in the *M. tuberculosis* genome, including the *Rv3616c–3612c* cluster, which encodes at least two ESX-1-secreted proteins and is required for the functioning of ESX-1¹¹ (Fig. 1).

Although it is not surprising that mycobacteria express a transcription factor that regulates the expression of ESX-1 components, that EspR is itself secreted by ESX-1 is an unexpected result. The authors propose that secretion of this protein constitutes an unusual feedback loop that could be part of a finely tuned control process to prevent excessive or prolonged activity of ESX-1 during infection of mammalian cells. Although secretion of a regulatory protein as a mechanism for diminishing its activity inside the cell has been described before, EspR might represent the first example of a transcription factor that is actively exported from the cell by the same secretion system that it induces.

One can imagine how EspR could impose a limit on the level and duration of ESX-1 activity. When initially expressed, this protein could be essential for activating ESX-1 to secrete high levels of virulence-promoting proteins, thus allowing *M. tuberculosis* to establish its infection in the host. Once infection is achieved, secretion of EspR might lead to a reduction of its transcriptional activity within the bacterium and diminished ESX-1 activity. This would partially attenuate virulence, thus favouring either bacterial persistence or a slow, chronic infection in order to enhance transmission.

Can EspR production be turned on or off, and — if so — what external stimuli and bacterial sensing and signalling molecules could be responsible for this? Are other components of the ESX-1 system separately regulated by factors distinct from EspR? These questions, together with the identification of other secreted ESX-1 substrates and their mechanisms of action in mammalian host cells, should provide many opportunities for

deciphering the unique logic of mycobacterial virulence strategies.

It will also be interesting to determine whether EspR has any specific function once it has been exported from the bacterial cell. Because at least two other ESX-1 secreted products — ESAT-6 and CFP-10 — are major targets for the immune response, it is possible that EspR is also a prominent mycobacterial antigen. If so, this could have implications for the development of new vaccines and diagnostic tools for tuberculosis.

Steven A. Porcelli is in the Departments of Microbiology and Immunology, and of Medicine, Albert Einstein College of Medicine, 1300 Morris Park Avenue, Bronx, New York 10461, USA. e-mail: porcelli@aecom.yu.edu

1. Dye, C., Scheele, S., Dolin, P., Pathania, V. & Ravignione, M. C. *J. Am. Med. Assoc.* **282**, 677–686 (1999).
2. Raghavan, S., Manzanillo, P., Chan, K., Dovey, C. & Cox, J. S. *Nature* **454**, 717–721 (2008).
3. Sorensen, A. L., Nagai, S., Houen, G., Andersen, P. & Andersen, A. B. *Infect. Immun.* **63**, 1710–1717 (1995).
4. Gey van Pittius, N. C. et al. *Genome Biol.* doi:10.1186/gb-2001-2-10-research0044 (2001).
5. Converse, S. E. & Cox, J. S. *J. Bacteriol.* **187**, 1238–1245 (2005).
6. DiGiuseppe Champion, P. A. & Cox, J. S. *Cell Microbiol.* **9**, 1376–1384 (2007).
7. Stanley, S. A., Raghavan, S., Hwang, W. W. & Cox, J. S. *Proc. Natl Acad. Sci. USA* **100**, 13001–13006 (2003).
8. van der Wel, N. et al. *Cell* **129**, 1287–1298 (2007).
9. Hsu, T. et al. *Proc. Natl Acad. Sci. USA* **100**, 12420–12425 (2003).
10. MacGurn, J. A. & Cox, J. S. *Infect. Immun.* **75**, 2668–2678 (2007).
11. MacGurn, J. A., Raghavan, S., Stanley, S. A. & Cox, J. S. *Mol. Microbiol.* **57**, 1653–1663 (2005).

OPTICS

Electronic eyeballs

Takao Someya

The ability to fabricate silicon optoelectronic devices on a curved surface will lead to imaging systems with exceptional characteristics. This innovative technology will find diverse applications.

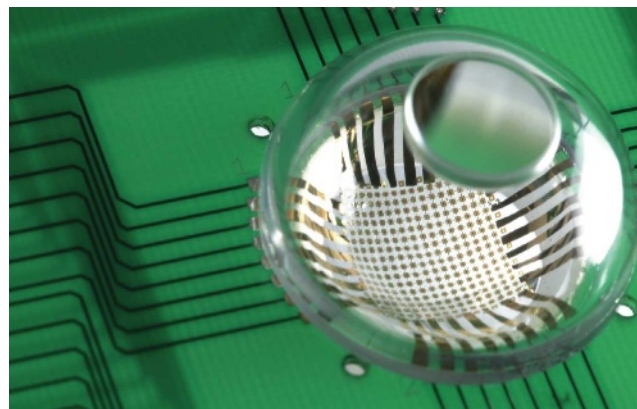
All animals have curved imagers for their eyes. By contrast, artificial vision systems such as digital and video cameras have to rely on flat image-recording surfaces. These artificial imagers are made with silicon microfabrication technologies to produce the necessary network of semiconductor photodetectors, and they can now create pictures with more than ten million pixels. But there remains the big problem of producing bright, distortion-free images with a flat imager. Given the distortion that occurs at the edges of lenses, multiple combinations of different lenses are required for effective imaging. Consequently, lens arrangements are heavy, expensive and produce darker results than they would otherwise do¹.

On page 748 of this issue, Rogers and colleagues (Ko et al.)² describe how they have drawn inspiration from animals' eyes and have

succeeded in eliminating these fundamental limitations of conventional artificial-vision systems. Their electronic eye camera (Fig. 1) is based on silicon electronics that is designed to have full mechanical compressibility–stretchability, meaning that it can be moulded onto a hemispherical substrate.

The authors' method is outlined in Figure 1 of their paper (page 749). It depends on two main advances. The first is the fabrication, on a silicon wafer, of a network of semiconductor photodetectors that can tolerate elastic compressibility despite being subjected to high levels of strain (typically exceeding 50%). The crucial features that allow such compressibility are the thin metallic wires that interconnect the photodetectors. The second innovation is the use of elastomeric elements that can transform a photodetector network initially made in a planar configuration into hemispherical

Figure 1 | The electronic eye camera. This device shows the integration of the concave photodetector system devised by Ko et al.² into a miniature camera that has a single, simple lens. Apart from the lens at the top, the hemispherical cap would not normally be transparent. The camera is about 2 centimetres in diameter. (Photo courtesy of J. A. Rogers.)



hinders phagosome maturation in macrophages through a process that does not require ESAT-6 or other known substrates secreted by this system¹⁰. Taken together, these data indicate that ESX-1 modulates innate immune responses of the infected host through several mechanisms, which probably involve ESX-1-mediated secretion of ESAT-6 and other unidentified factors.

Raghavan and colleagues² identify a previously unknown component of ESX-1 that is not only a central regulator of it but is also secreted by this system. They find that, in *M. tuberculosis*, the *Rv3849* gene — which is located some distance from the main ESX-1 gene cluster surrounding the *esxB* operon — is required for ESX-1 function. The protein product of *Rv3849*, EspR, is highly similar to a gene transcription factor of the harmless soil bacterium *Bacillus subtilis*. The authors find that EspR is also a DNA-binding transcriptional regulator — *Rv3849* deactivation leads to changes in the transcription of a few operons in the *M. tuberculosis* genome, including the *Rv3616c–3612c* cluster, which encodes at least two ESX-1-secreted proteins and is required for the functioning of ESX-1¹¹ (Fig. 1).

Although it is not surprising that mycobacteria express a transcription factor that regulates the expression of ESX-1 components, that EspR is itself secreted by ESX-1 is an unexpected result. The authors propose that secretion of this protein constitutes an unusual feedback loop that could be part of a finely tuned control process to prevent excessive or prolonged activity of ESX-1 during infection of mammalian cells. Although secretion of a regulatory protein as a mechanism for diminishing its activity inside the cell has been described before, EspR might represent the first example of a transcription factor that is actively exported from the cell by the same secretion system that it induces.

One can imagine how EspR could impose a limit on the level and duration of ESX-1 activity. When initially expressed, this protein could be essential for activating ESX-1 to secrete high levels of virulence-promoting proteins, thus allowing *M. tuberculosis* to establish its infection in the host. Once infection is achieved, secretion of EspR might lead to a reduction of its transcriptional activity within the bacterium and diminished ESX-1 activity. This would partially attenuate virulence, thus favouring either bacterial persistence or a slow, chronic infection in order to enhance transmission.

Can EspR production be turned on or off, and — if so — what external stimuli and bacterial sensing and signalling molecules could be responsible for this? Are other components of the ESX-1 system separately regulated by factors distinct from EspR? These questions, together with the identification of other secreted ESX-1 substrates and their mechanisms of action in mammalian host cells, should provide many opportunities for

deciphering the unique logic of mycobacterial virulence strategies.

It will also be interesting to determine whether EspR has any specific function once it has been exported from the bacterial cell. Because at least two other ESX-1 secreted products — ESAT-6 and CFP-10 — are major targets for the immune response, it is possible that EspR is also a prominent mycobacterial antigen. If so, this could have implications for the development of new vaccines and diagnostic tools for tuberculosis.

Steven A. Porcelli is in the Departments of Microbiology and Immunology, and of Medicine, Albert Einstein College of Medicine, 1300 Morris Park Avenue, Bronx, New York 10461, USA. e-mail: porcelli@aecom.yu.edu

1. Dye, C., Scheele, S., Dolin, P., Pathania, V. & Ravignione, M. C. *J. Am. Med. Assoc.* **282**, 677–686 (1999).
2. Raghavan, S., Manzanillo, P., Chan, K., Dovey, C. & Cox, J. S. *Nature* **454**, 717–721 (2008).
3. Sorensen, A. L., Nagai, S., Houen, G., Andersen, P. & Andersen, A. B. *Infect. Immun.* **63**, 1710–1717 (1995).
4. Gey van Pittius, N. C. et al. *Genome Biol.* doi:10.1186/gb-2001-2-10-research0044 (2001).
5. Converse, S. E. & Cox, J. S. *J. Bacteriol.* **187**, 1238–1245 (2005).
6. DiGiuseppe Champion, P. A. & Cox, J. S. *Cell Microbiol.* **9**, 1376–1384 (2007).
7. Stanley, S. A., Raghavan, S., Hwang, W. W. & Cox, J. S. *Proc. Natl Acad. Sci. USA* **100**, 13001–13006 (2003).
8. van der Wel, N. et al. *Cell* **129**, 1287–1298 (2007).
9. Hsu, T. et al. *Proc. Natl Acad. Sci. USA* **100**, 12420–12425 (2003).
10. MacGurn, J. A. & Cox, J. S. *Infect. Immun.* **75**, 2668–2678 (2007).
11. MacGurn, J. A., Raghavan, S., Stanley, S. A. & Cox, J. S. *Mol. Microbiol.* **57**, 1653–1663 (2005).

OPTICS

Electronic eyeballs

Takao Someya

The ability to fabricate silicon optoelectronic devices on a curved surface will lead to imaging systems with exceptional characteristics. This innovative technology will find diverse applications.

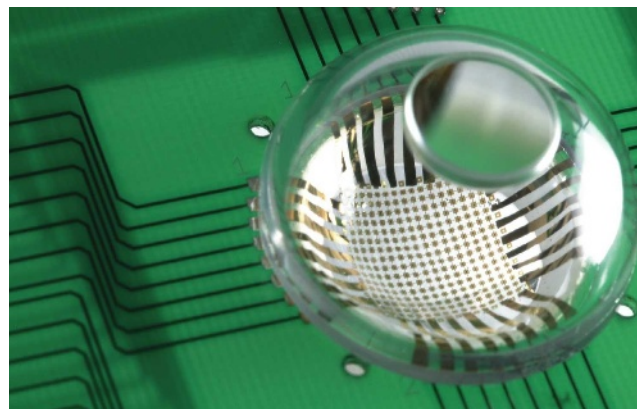
All animals have curved imagers for their eyes. By contrast, artificial vision systems such as digital and video cameras have to rely on flat image-recording surfaces. These artificial imagers are made with silicon microfabrication technologies to produce the necessary network of semiconductor photodetectors, and they can now create pictures with more than ten million pixels. But there remains the big problem of producing bright, distortion-free images with a flat imager. Given the distortion that occurs at the edges of lenses, multiple combinations of different lenses are required for effective imaging. Consequently, lens arrangements are heavy, expensive and produce darker results than they would otherwise do¹.

On page 748 of this issue, Rogers and colleagues (Ko et al.)² describe how they have drawn inspiration from animals' eyes and have

succeeded in eliminating these fundamental limitations of conventional artificial-vision systems. Their electronic eye camera (Fig. 1) is based on silicon electronics that is designed to have full mechanical compressibility–stretchability, meaning that it can be moulded onto a hemispherical substrate.

The authors' method is outlined in Figure 1 of their paper (page 749). It depends on two main advances. The first is the fabrication, on a silicon wafer, of a network of semiconductor photodetectors that can tolerate elastic compressibility despite being subjected to high levels of strain (typically exceeding 50%). The crucial features that allow such compressibility are the thin metallic wires that interconnect the photodetectors. The second innovation is the use of elastomeric elements that can transform a photodetector network initially made in a planar configuration into hemispherical

Figure 1 | The electronic eye camera. This device shows the integration of the concave photodetector system devised by Ko et al.² into a miniature camera that has a single, simple lens. Apart from the lens at the top, the hemispherical cap would not normally be transparent. The camera is about 2 centimetres in diameter. (Photo courtesy of J. A. Rogers.)



shapes for implementation in imagers. This technology heralds the advent of new classes of imaging devices with wide-angle fields of view, low distortion and compact size.

The recent developments³ in compressible–stretchable electronics provide the prospect of many new applications such as long-term bionic implants, robotic sensory skins⁴, ambient displays embedded in (for example) wallpaper, and intelligent surfaces that are chemically or electronically functionalized and can interact with people, objects or their environment. One of the most difficult requirements is to achieve excellent mechanical robustness and good electronic performance while satisfying basic electrical requirements — the materials and circuit architecture used in conformable and stretchable electronics must be designed such that their mechanical integrity and electrical functionality are preserved during the fabrication and use of the resulting products. In recent years, Rogers and his colleagues³ have developed one- and two-dimensional stretchable ribbons and circuits for this purpose; the two-dimensional compressible components used in the electronic eye camera represent a natural extension of this line of research.

The promise of this technology extends well beyond the hemispheric configuration demonstrated by Ko *et al.*². For instance, it could be applied to integrate optoelectronics onto complex, curvilinear surfaces for use in health-monitoring devices that optically detect concentrations of oxygen and other constituents in blood. The new possibilities in optics design should lead to a further reduction in the imaging distortion of ultra-compact camera systems in which photodetector surface geometries can be carefully optimized. Furthermore, distortion-free, adaptive focusing mechanisms might be feasible if the stretchable imager of these camera systems can be developed on actively deformable substrate surfaces. Such simplified systems should have much improved optical transparency — that is, have much reduced optical loss compared with that arising from the use of multiple lenses. This beneficial feature will not only generate more industrial applications for these systems, it will also benefit fundamental research at wavelengths for which existing materials cannot ensure sufficient optical transparency.

In addition to the further development of concave photodetector systems, we can expect to see advances in creating convex imagers — for use in, for example, artificial insect-like compound eyes with exceptional dynamic visual acuity, and in fish eyes that have a 360° field of view. These and other types of biologically inspired device should become feasible given the advances in optical engineering made possible by the advent of geometrically transformable and stretchable–compressible electronics and optoelectronics. All in all, with their electronic eye camera, Rogers and colleagues have delivered an outstanding contribution by showing how progress in

electronics can be made by overcoming the constraints of flat silicon wafers.

Takao Someya is in the Quantum-Phase Electronics Center, School of Engineering, the University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan.

e-mail: someya@ap.t.u-tokyo.ac.jp

1. Born, M. & Wolf, E. *Principles of Optics* (Pergamon, 1959).
2. Ko, H. C. *et al.* *Nature* **454**, 748–753 (2008).
3. Lacour, S. P., Jones, J., Wagner, S., Li, T. & Suo, Z. *Proc. IEEE* **93**, 1459–1467 (2005).
4. Someya, T. *et al.* *Proc. Natl Acad. Sci. USA* **102**, 12321–12325 (2005).
5. Choi, W. M. *et al.* *Nano Lett.* **7**, 1655–1663 (2007).

PHARMACOLOGY

Unready for action

Joe Henry Steinbach

Boy scouts recognize that the key to success is to be prepared. The same is true of molecules that bind to and open ion channels — the least effective ones are slower to prepare the channel to be ready for opening.

Some drugs, known collectively as agonists, can be thought of as molecular switches — if the molecule fits the active site of the receptor, the biological response is switched on. But so-called partial agonists pose a problem for this simple model. A partial agonist is a compound that elicits less than a full biological response on binding to its target, even when it occupies all the available binding sites. How can this be?

Reporting on page 722 of this issue, Lape *et al.*¹ provide an answer for partial agonists that bind to two structurally related channels — the glycine receptor and the muscle nicotinic receptor. The binding of a full agonist to these receptors causes the channels to be open almost constantly, so that the maximum possible current in the channel is observed as ions flow through. But when partial agonists bind, the channel is open for a smaller proportion of the time, and only a fraction of the maximum current flows. Lape *et al.* show that this is because the partial agonists often fail to trigger a conformational change in the receptor that precedes the actual opening of the channel.

The classic view of drug action is known as the occupancy model, and proposes that, when a drug binds to an effector (a receptor or an enzyme), the drug–effector complex constitutes the signal that generates a biological effect (Fig. 1a). The discovery of partial agonists posed a problem for this simple model, so an additional step was added²: binding creates an inert drug–effector complex, which then undergoes a conformational change to yield an active state. In this scheme, partial agonists binding to ion-channel receptors were thought to cause those channels to open slowly (Fig. 1b). This idea has dominated thinking about the nature of partial agonism for the past 50 years.

But a third model for receptor activation has also been proposed, in which an intermediate state exists between the initial, inert drug–effector complex and the receptor with the channel open (Fig. 1c). This intervening state has been called the flip state³. Readers of

a military bent might prefer to think of it as a cocked state.

The flip-state theory certainly makes sense for ion channels. These large proteins consist of several subunits, and their activation involves a considerable conformational change that probably takes place in a series of steps. Thermodynamic analyses of kinetic experiments have been used to infer the relative times at which individual amino acids in muscle nicotinic acetylcholine receptors are perturbed during activation^{4,5}. The results suggest that five distinct sets of residues exist, with those near the agonist binding site moving first, and those near the channel's gate moving later. This kind of analysis has also been used to compare a range of molecules that bind to acetylcholine receptors, from weak partial agonists to full agonists⁶. The differences between agonists appeared near the start of the activation process, at the same time as the movements of amino acids close to the agonist binding site.

The best way to study receptor states as agonists bind is by the kinetic analysis of currents through a single ion channel. Lape *et al.*¹ adopted this approach, analysing the actions of partial agonists using high-resolution, single-channel recordings. They examined the durations of the brief periods in which a channel is closed while an agonist or a partial agonist is bound, and confirm the existence of a flip state. Their results also show why partial agonists fail to maximally activate these ion channels. It seems that partial agonists do not have an intrinsically low channel-opening rate. In fact, they are just as good at opening channels as full agonists — that is, the rate of conversion of the flip state to the open state is as high as for full agonists. The difference is that partial agonists are ineffective at converting the inert drug–receptor complex to the flip state, so their overall ability to produce open channels is low.

It would be grand to know what the flip state looks like. Crystallography can provide satisfying pictures of proteins, although it can be difficult to relate the resulting static images to

shapes for implementation in imagers. This technology heralds the advent of new classes of imaging devices with wide-angle fields of view, low distortion and compact size.

The recent developments³ in compressible–stretchable electronics provide the prospect of many new applications such as long-term bionic implants, robotic sensory skins⁴, ambient displays embedded in (for example) wallpaper, and intelligent surfaces that are chemically or electronically functionalized and can interact with people, objects or their environment. One of the most difficult requirements is to achieve excellent mechanical robustness and good electronic performance while satisfying basic electrical requirements — the materials and circuit architecture used in conformable and stretchable electronics must be designed such that their mechanical integrity and electrical functionality are preserved during the fabrication and use of the resulting products. In recent years, Rogers and his colleagues³ have developed one- and two-dimensional stretchable ribbons and circuits for this purpose; the two-dimensional compressible components used in the electronic eye camera represent a natural extension of this line of research.

The promise of this technology extends well beyond the hemispheric configuration demonstrated by Ko *et al.*². For instance, it could be applied to integrate optoelectronics onto complex, curvilinear surfaces for use in health-monitoring devices that optically detect concentrations of oxygen and other constituents in blood. The new possibilities in optics design should lead to a further reduction in the imaging distortion of ultra-compact camera systems in which photodetector surface geometries can be carefully optimized. Furthermore, distortion-free, adaptive focusing mechanisms might be feasible if the stretchable imager of these camera systems can be developed on actively deformable substrate surfaces. Such simplified systems should have much improved optical transparency — that is, have much reduced optical loss compared with that arising from the use of multiple lenses. This beneficial feature will not only generate more industrial applications for these systems, it will also benefit fundamental research at wavelengths for which existing materials cannot ensure sufficient optical transparency.

In addition to the further development of concave photodetector systems, we can expect to see advances in creating convex imagers — for use in, for example, artificial insect-like compound eyes with exceptional dynamic visual acuity, and in fish eyes that have a 360° field of view. These and other types of biologically inspired device should become feasible given the advances in optical engineering made possible by the advent of geometrically transformable and stretchable–compressible electronics and optoelectronics. All in all, with their electronic eye camera, Rogers and colleagues have delivered an outstanding contribution by showing how progress in

electronics can be made by overcoming the constraints of flat silicon wafers.

Takao Someya is in the Quantum-Phase Electronics Center, School of Engineering, the University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan.

e-mail: someya@ap.t.u-tokyo.ac.jp

1. Born, M. & Wolf, E. *Principles of Optics* (Pergamon, 1959).
2. Ko, H. C. *et al. Nature* **454**, 748–753 (2008).
3. Lacour, S. P., Jones, J., Wagner, S., Li, T. & Suo, Z. *Proc. IEEE* **93**, 1459–1467 (2005).
4. Someya, T. *et al. Proc. Natl Acad. Sci. USA* **102**, 12321–12325 (2005).
5. Choi, W. M. *et al. Nano Lett.* **7**, 1655–1663 (2007).

PHARMACOLOGY

Unready for action

Joe Henry Steinbach

Boy scouts recognize that the key to success is to be prepared. The same is true of molecules that bind to and open ion channels — the least effective ones are slower to prepare the channel to be ready for opening.

Some drugs, known collectively as agonists, can be thought of as molecular switches — if the molecule fits the active site of the receptor, the biological response is switched on. But so-called partial agonists pose a problem for this simple model. A partial agonist is a compound that elicits less than a full biological response on binding to its target, even when it occupies all the available binding sites. How can this be?

Reporting on page 722 of this issue, Lape *et al.*¹ provide an answer for partial agonists that bind to two structurally related channels — the glycine receptor and the muscle nicotinic receptor. The binding of a full agonist to these receptors causes the channels to be open almost constantly, so that the maximum possible current in the channel is observed as ions flow through. But when partial agonists bind, the channel is open for a smaller proportion of the time, and only a fraction of the maximum current flows. Lape *et al.* show that this is because the partial agonists often fail to trigger a conformational change in the receptor that precedes the actual opening of the channel.

The classic view of drug action is known as the occupancy model, and proposes that, when a drug binds to an effector (a receptor or an enzyme), the drug–effector complex constitutes the signal that generates a biological effect (Fig. 1a). The discovery of partial agonists posed a problem for this simple model, so an additional step was added²: binding creates an inert drug–effector complex, which then undergoes a conformational change to yield an active state. In this scheme, partial agonists binding to ion-channel receptors were thought to cause those channels to open slowly (Fig. 1b). This idea has dominated thinking about the nature of partial agonism for the past 50 years.

But a third model for receptor activation has also been proposed, in which an intermediate state exists between the initial, inert drug–effector complex and the receptor with the channel open (Fig. 1c). This intervening state has been called the flip state³. Readers of

a military bent might prefer to think of it as a cocked state.

The flip-state theory certainly makes sense for ion channels. These large proteins consist of several subunits, and their activation involves a considerable conformational change that probably takes place in a series of steps. Thermodynamic analyses of kinetic experiments have been used to infer the relative times at which individual amino acids in muscle nicotinic acetylcholine receptors are perturbed during activation^{4,5}. The results suggest that five distinct sets of residues exist, with those near the agonist binding site moving first, and those near the channel's gate moving later. This kind of analysis has also been used to compare a range of molecules that bind to acetylcholine receptors, from weak partial agonists to full agonists⁶. The differences between agonists appeared near the start of the activation process, at the same time as the movements of amino acids close to the agonist binding site.

The best way to study receptor states as agonists bind is by the kinetic analysis of currents through a single ion channel. Lape *et al.*¹ adopted this approach, analysing the actions of partial agonists using high-resolution, single-channel recordings. They examined the durations of the brief periods in which a channel is closed while an agonist or a partial agonist is bound, and confirm the existence of a flip state. Their results also show why partial agonists fail to maximally activate these ion channels. It seems that partial agonists do not have an intrinsically low channel-opening rate. In fact, they are just as good at opening channels as full agonists — that is, the rate of conversion of the flip state to the open state is as high as for full agonists. The difference is that partial agonists are ineffective at converting the inert drug–receptor complex to the flip state, so their overall ability to produce open channels is low.

It would be grand to know what the flip state looks like. Crystallography can provide satisfying pictures of proteins, although it can be difficult to relate the resulting static images to

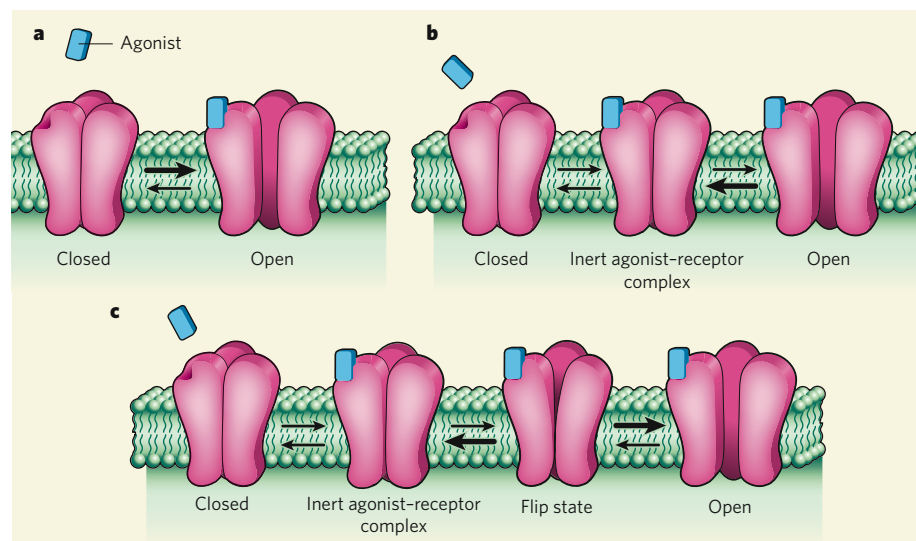


Figure 1 | Models for partial agonism. Agonists bind to receptors to produce an effect, such as the opening of an ion channel. **a**, The occupancy model proposes that agonist binding opens the channel immediately. But this model does not explain how a so-called partial agonist can bind to produce a complex in which the probability of channel opening is low. **b**, A two-step mechanism has therefore been proposed² in which an inert receptor–agonist complex forms first, after which the channel opens. Partial agonists were thought to be less effective than full agonists in the channel-opening step. **c**, Lape *et al.*¹ show that another step is required, in which the inert complex enters a state of readiness — the flip state — before channel opening. Partial agonists are less effective than full agonists at inducing the flip state, but bring about channel opening just as quickly.

functional states. Unfortunately, no complete high-resolution structures have been determined for these receptors in known functional states. The crystal structures of putative flip states will be even more difficult to obtain, as

such states are necessarily short-lived so that the protein can respond rapidly to transient changes in agonist concentration. Optical or resonance techniques (such as nuclear magnetic resonance) might fill this gap in the future.

The authors' insight¹ into the mechanistic basis of partial agonism will quell some angst among pharmacologists, but it also generates ideas about other consequences of the flip state. For example, it has been reported that mutations in the glycine receptors of mice reduce the ability of glycine to open the associated ion channels, because the receptor is less able to 'flip' after binding the agonist⁷. It is also known that some compounds enhance the effects of partial agonists, and can even sometimes convert them to full agonists. Perhaps these compounds affect the entry of the channel to the flip state. Similarly, certain inhibitors of agonists might work by reducing the stability of the flip state.

Lape *et al.*¹ have answered a long-standing question by showing that the flip state can explain partial agonism. Future research will reveal what other aspects of receptor function are shaped by the properties of this state. ■

Joe Henry Steinbach is in the Department of Anesthesiology, Washington University School of Medicine, St Louis, Missouri 63110, USA. e-mail: jhs@morpheus.wustl.edu

1. Lape, R., Colquhoun, D. & Sivillotti, L. G. *Nature* **454**, 722–727 (2008).
2. del Castillo, J. & Katz, B. *Proc. R. Soc. Lond. B Biol. Sci.* **146**, 362–368 (1957).
3. Burzomato, V. *et al. J. Neurosci.* **24**, 10924–10940 (2004).
4. Bafna, P. A., Purohit, P. G. & Auerbach, A. *PLoS ONE* **3**, e2515 (2008).
5. Auerbach, A. *Proc. Natl Acad. Sci. USA* **102**, 1408–1412 (2005).
6. Grosman, C., Zhou, M. & Auerbach, A. *Nature* **403**, 773–776 (2000).
7. Plested, A. J. R. *et al. J. Physiol. (Lond.)* **581**, 51–73 (2007).

MATERIALS SCIENCE

Protein gels on the move

April M. Kloxin and Kristi S. Anseth

Light-induced reactions enable three-dimensional objects to be built from simple compounds. Proteins have been added to the list of building blocks, and the resulting gels move in response to environmental cues.

The ancient Egyptians were arguably among the first photochemists, because they used light-induced reactions as part of their mummification process. They wrapped bodies in linens that had been dipped in a substance called bitumen of Judea. On exposure to sunlight, the molecules in the bitumen reacted to form crosslinks with each other, yielding a hard, protective coating that preserved the mummy for centuries¹. Variations of this crude polymerization process are still in use today. Latter-day scientists can choose from thousands of different light-reactive monomers, and use sophisticated laser techniques to turn them into polymers that have diverse applications — for forming printing plates for newspapers, for example, or for dental restoration.

Reporting in *Proceedings of the National Academy of Sciences*, Kaehr and Shear² describe

a fresh twist to this technique. They have generated complex materials by using lasers to induce crosslinking reactions in proteins. Remarkably, the resulting three-dimensional objects change shape in response to their environment. Such systems might find use as moving parts in miniaturized devices, with applications in medical diagnostics or remote environmental sensing.

The authors' approach relies on advances in a technique known as three-dimensional light patterning. Traditionally, this involves 'drawing' a pattern with a narrow laser beam across liquid monomers on a surface. The monomers react to form solid polymers at the places where the laser strikes, so that three-dimensional objects can be built up one layer at a time. A faster process is also now used, in which an entire layer of polymer is made at once by passing

light through a stencil — known as a mask to those in the trade — of the desired pattern.

Kaehr and Shear combine these two approaches by raster-scanning a narrow laser beam across a mask on a protein solution, to create solid, three-dimensional structures. This allows complex macroscopic structures to be made without the elaborate and expensive apparatus required to raster-scan precise shapes with a laser. Using their technique, the authors can expeditiously generate objects with dimensions as small as 100 nanometres, and of any desired shape.

To date, proteins have not found much use as building blocks for three-dimensional photofabrication processes, because commonly used photochemical reactions typically denature or destroy the molecules. An unexpected breakthrough came from studies of photodynamic therapy, a medical procedure that uses molecules called photosensitizers to absorb light and so destroy tumours. The mechanism of action was thought to involve photochemical reactions that created covalent links between proteins (or between proteins and other molecules) in tumour cells, leading to cell damage. The crosslinking was generally considered to be uncontrolled, occurring randomly between the various kinds of amino acid. But studies on proteins³, and on synthetic macromolecules

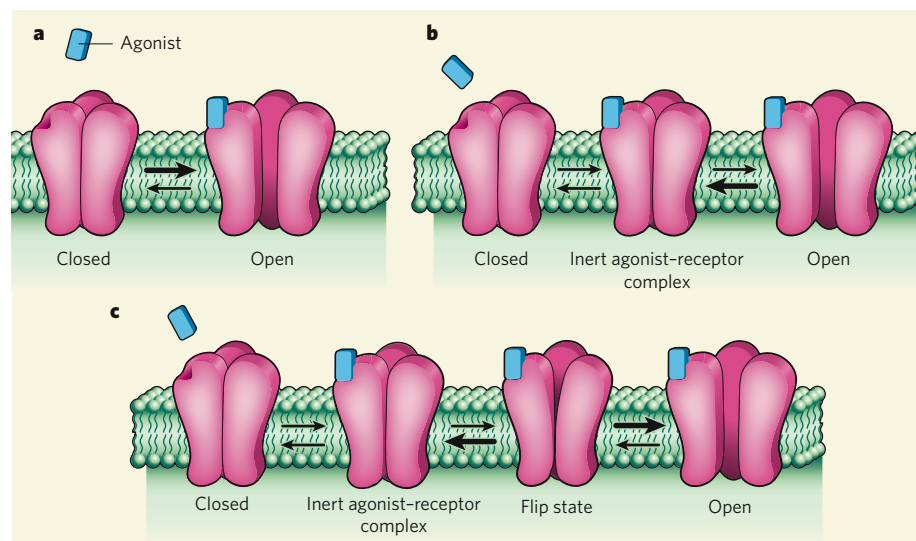


Figure 1 | Models for partial agonism. Agonists bind to receptors to produce an effect, such as the opening of an ion channel. **a**, The occupancy model proposes that agonist binding opens the channel immediately. But this model does not explain how a so-called partial agonist can bind to produce a complex in which the probability of channel opening is low. **b**, A two-step mechanism has therefore been proposed² in which an inert receptor–agonist complex forms first, after which the channel opens. Partial agonists were thought to be less effective than full agonists in the channel-opening step. **c**, Lape *et al.*¹ show that another step is required, in which the inert complex enters a state of readiness — the flip state — before channel opening. Partial agonists are less effective than full agonists at inducing the flip state, but bring about channel opening just as quickly.

functional states. Unfortunately, no complete high-resolution structures have been determined for these receptors in known functional states. The crystal structures of putative flip states will be even more difficult to obtain, as

such states are necessarily short-lived so that the protein can respond rapidly to transient changes in agonist concentration. Optical or resonance techniques (such as nuclear magnetic resonance) might fill this gap in the future.

The authors' insight¹ into the mechanistic basis of partial agonism will quell some angst among pharmacologists, but it also generates ideas about other consequences of the flip state. For example, it has been reported that mutations in the glycine receptors of mice reduce the ability of glycine to open the associated ion channels, because the receptor is less able to 'flip' after binding the agonist⁷. It is also known that some compounds enhance the effects of partial agonists, and can even sometimes convert them to full agonists. Perhaps these compounds affect the entry of the channel to the flip state. Similarly, certain inhibitors of agonists might work by reducing the stability of the flip state.

Lape *et al.*¹ have answered a long-standing question by showing that the flip state can explain partial agonism. Future research will reveal what other aspects of receptor function are shaped by the properties of this state. ■

Joe Henry Steinbach is in the Department of Anesthesiology, Washington University School of Medicine, St Louis, Missouri 63110, USA. e-mail: jhs@morpheus.wustl.edu

1. Lape, R., Colquhoun, D. & Sivillotti, L. G. *Nature* **454**, 722–727 (2008).
2. del Castillo, J. & Katz, B. *Proc. R. Soc. Lond. B Biol. Sci.* **146**, 362–368 (1957).
3. Burzomato, V. *et al. J. Neurosci.* **24**, 10924–10940 (2004).
4. Bafna, P. A., Purohit, P. G. & Auerbach, A. *PLoS ONE* **3**, e2515 (2008).
5. Auerbach, A. *Proc. Natl Acad. Sci. USA* **102**, 1408–1412 (2005).
6. Grosman, C., Zhou, M. & Auerbach, A. *Nature* **403**, 773–776 (2000).
7. Plested, A. J. R. *et al. J. Physiol. (Lond.)* **581**, 51–73 (2007).

MATERIALS SCIENCE

Protein gels on the move

April M. Kloxin and Kristi S. Anseth

Light-induced reactions enable three-dimensional objects to be built from simple compounds. Proteins have been added to the list of building blocks, and the resulting gels move in response to environmental cues.

The ancient Egyptians were arguably among the first photochemists, because they used light-induced reactions as part of their mummification process. They wrapped bodies in linens that had been dipped in a substance called bitumen of Judea. On exposure to sunlight, the molecules in the bitumen reacted to form crosslinks with each other, yielding a hard, protective coating that preserved the mummy for centuries¹. Variations of this crude polymerization process are still in use today. Latter-day scientists can choose from thousands of different light-reactive monomers, and use sophisticated laser techniques to turn them into polymers that have diverse applications — for forming printing plates for newspapers, for example, or for dental restoration.

Reporting in *Proceedings of the National Academy of Sciences*, Kaehr and Shear² describe

a fresh twist to this technique. They have generated complex materials by using lasers to induce crosslinking reactions in proteins. Remarkably, the resulting three-dimensional objects change shape in response to their environment. Such systems might find use as moving parts in miniaturized devices, with applications in medical diagnostics or remote environmental sensing.

The authors' approach relies on advances in a technique known as three-dimensional light patterning. Traditionally, this involves 'drawing' a pattern with a narrow laser beam across liquid monomers on a surface. The monomers react to form solid polymers at the places where the laser strikes, so that three-dimensional objects can be built up one layer at a time. A faster process is also now used, in which an entire layer of polymer is made at once by passing

light through a stencil — known as a mask to those in the trade — of the desired pattern.

Kaehr and Shear combine these two approaches by raster-scanning a narrow laser beam across a mask on a protein solution, to create solid, three-dimensional structures. This allows complex macroscopic structures to be made without the elaborate and expensive apparatus required to raster-scan precise shapes with a laser. Using their technique, the authors can expeditiously generate objects with dimensions as small as 100 nanometres, and of any desired shape.

To date, proteins have not found much use as building blocks for three-dimensional photo-fabrication processes, because commonly used photochemical reactions typically denature or destroy the molecules. An unexpected breakthrough came from studies of photodynamic therapy, a medical procedure that uses molecules called photosensitizers to absorb light and so destroy tumours. The mechanism of action was thought to involve photochemical reactions that created covalent links between proteins (or between proteins and other molecules) in tumour cells, leading to cell damage. The crosslinking was generally considered to be uncontrolled, occurring randomly between the various kinds of amino acid. But studies on proteins³, and on synthetic macromolecules

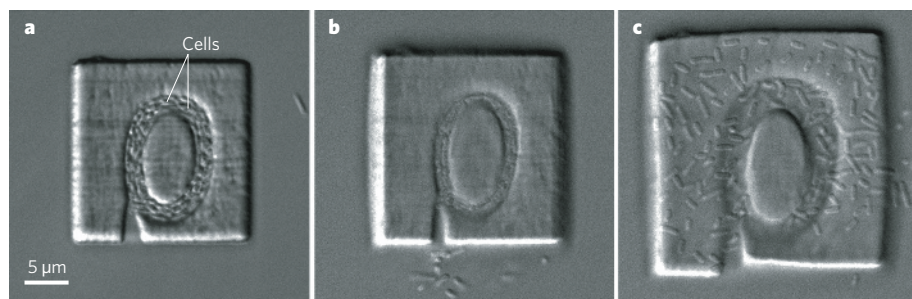


Figure 1 | Dynamic microchambers for cell culture. Kaehr and Shear² have prepared gels that change shape in response to environmental cues, for example pH. The materials can be made into components for various devices, such as chambers for culturing cells. **a**, This doughnut-shaped chamber, in a bath at pH 7, contains cultured *Escherichia coli* cells. **b**, A change of acidity (to pH 12.2) causes compression of the chamber and the release of a few cells. **c**, Eventually, the chamber is completely disrupted, and all the cells escape into the surrounding environment.

containing selected amino-acid side chains⁴, conclusively demonstrated that crosslinking of only specific amino acids occurs in the presence of a photosensitizer on exposure to light.

Kaehr and Shear² exploited this controlled crosslinking process to make protein-based materials that have highly defined shapes and microstructures. Using their patterning technique in conjunction with a photosensitizer, they 'wrote' three-dimensional structures by crosslinking proteins in solution using spatially directed light. The entire process takes only a few minutes, and the resulting materials expand and contract, open and close, or flex in response to changes in their environment (such as changes in pH) — just as naturally occurring proteins do.

The rich chemistry of proteins, and the ease with which their properties can be modified by changing their amino-acid sequences, make this a versatile approach for synthesizing designer components for devices. Proteins can respond to several triggers, such as external changes in temperature or ionic concentration. It is also possible to control the number of molecular interactions between different proteins, to integrate several proteins into a single material, and to make three-dimensional protein structures at different scales. And, unlike synthetic macromolecules (polymers), proteins can be used that have uniform size, amino-acid sequence and three-dimensional structure. The combination of functional flexibility and the ability to make identical copies of each structure opens up many potential applications for crosslinked proteins as device components.

The authors demonstrated an application of their materials by using them to make dynamic enclosures for culturing cells (Fig. 1). The enclosures trap cells, expand as the cells multiply, and subsequently release the cultured cells in response to a change in pH. In the future, the properties of the protein gels used to make the microchambers could be modified by forming hybrid materials with synthetic polymers, and by using proteins that respond to a specific chemical signal. This approach has recently been demonstrated⁵ in a similar material, in which a modified version of the protein

calmodulin was trapped within a polymer network. The volume of the resulting gel changed when calcium ions or ligand molecules bound to the protein.

So what does the future hold for Kaehr and Shear's sophisticated protein gels? As potentially useful biomaterials, opportunities abound. The sequence, concentration and spatial presentation of biological signals within the material — such as binding pockets or the active sites of enzymes — can easily be varied, and the crosslinked proteins can be designed so that an external stimulus exposes such embedded features. And by using enzymes as a basis

for the gels, materials could be designed that amplify chemical signals or generate multiple responses from a single input.

Nevertheless, it won't all be plain sailing. Many proteins denature in response to changes in temperature or other environmental cues, and it isn't known whether all crosslinked proteins will undergo reversible changes in response to such signals. But as the repertoire of proteins that can be used in these materials expands, a tantalizing array of prospects can be envisaged: customized gels with passive and active molecular barriers to sift or release soluble chemical signals; systems that combine several design elements to mimic the complexity of naturally occurring tissues; or even actuators that have dynamic responses reminiscent of muscle contraction. ■

April M. Kloxin and Kristi S. Anseth are in the Department of Chemical and Biological Engineering, University of Colorado, Boulder, Colorado 80309-0424, USA.

e-mails: april.kloxin@colorado.edu; kristi.anseth@colorado.edu

1. Decker, C. J. *Coatings Technol.* **59**, 97–106 (1987).
2. Kaehr, B. & Shear, J. B. *Proc. Natl Acad. Sci. USA* **105**, 8850–8854 (2008).
3. Shen, H.-R., Spikes, J. D., Kopečeková, P. & Kopeček, J. *J. Photochem. Photobiol. B* **35**, 213–219 (1996).
4. Shen, H.-R., Spikes, J. D., Kopečeková, P. & Kopeček, J. *J. Photochem. Photobiol. B* **34**, 203–210 (1996).
5. Murphy, W. L., Dillmore, W. S., Modica, J. & Mrksich, M. *Angew. Chem. Int. Edn* **46**, 3066–3069 (2007).

ALZHEIMER'S DISEASE

The latest suspect

Rudolph E. Tanzi and Lars Bertram

Many genetic mutations and several environmental factors contribute to Alzheimer's disease. Yet another disease risk gene, one that is involved in calcium regulation, has been added to the mix.

Alzheimer's disease is characterized by neurodegeneration in several brain areas, with the hippocampus being one of the first regions affected. Accumulation of aggregates of the amyloid- β peptide, for example in the form of senile plaques, is another feature of this disorder. Dreses-Werringloer *et al.*¹ set out to search for genes that are preferentially expressed in the hippocampus and lie in chromosomal regions that have been linked² to the risk of developing Alzheimer's disease. Writing in *Cell*, they report the identification of a new Alzheimer's disease risk gene that not only meets both of these criteria, but also seems to modulate extracellular levels of amyloid- β .

The authors call this gene *CALHM1* (for calcium homeostasis modulator 1), because its product seems to be involved in calcium homeostasis. They show that increased expression of the *CALHM1* protein, which they find

to be localized in the cell membrane and in the membrane of an intracellular organelle known as the endoplasmic reticulum, enhances the entry of calcium into the cell and thus its concentration in the cytoplasm. They find that this results in changes in the processing of the amyloid precursor protein that lead to reduced levels of amyloid- β peptide.

Intrigued by these observations, the authors resequenced *CALHM1* from a small number of patients with Alzheimer's disease and healthy controls. They identify a single nucleotide polymorphism (variation) in this gene, which, at the protein level, results in the substitution of the proline amino acid at residue 86 with a leucine residue (P86L). In normal subjects, the gene version carrying this polymorphism (the leucine allele) seems to occur with a frequency of about 20%. In the patients with Alzheimer's, however, it occurs nearly twice as frequently. The authors then tested this polymorphism

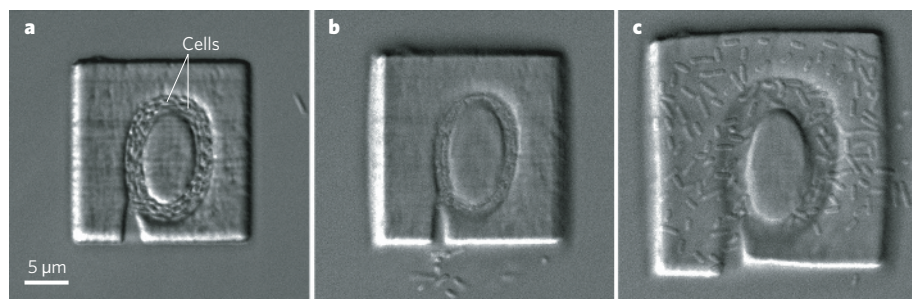


Figure 1 | Dynamic microchambers for cell culture. Kaehr and Shear² have prepared gels that change shape in response to environmental cues, for example pH. The materials can be made into components for various devices, such as chambers for culturing cells. **a**, This doughnut-shaped chamber, in a bath at pH 7, contains cultured *Escherichia coli* cells. **b**, A change of acidity (to pH 12.2) causes compression of the chamber and the release of a few cells. **c**, Eventually, the chamber is completely disrupted, and all the cells escape into the surrounding environment.

containing selected amino-acid side chains⁴, conclusively demonstrated that crosslinking of only specific amino acids occurs in the presence of a photosensitizer on exposure to light.

Kaehr and Shear² exploited this controlled crosslinking process to make protein-based materials that have highly defined shapes and microstructures. Using their patterning technique in conjunction with a photosensitizer, they 'wrote' three-dimensional structures by crosslinking proteins in solution using spatially directed light. The entire process takes only a few minutes, and the resulting materials expand and contract, open and close, or flex in response to changes in their environment (such as changes in pH) — just as naturally occurring proteins do.

The rich chemistry of proteins, and the ease with which their properties can be modified by changing their amino-acid sequences, make this a versatile approach for synthesizing designer components for devices. Proteins can respond to several triggers, such as external changes in temperature or ionic concentration. It is also possible to control the number of molecular interactions between different proteins, to integrate several proteins into a single material, and to make three-dimensional protein structures at different scales. And, unlike synthetic macromolecules (polymers), proteins can be used that have uniform size, amino-acid sequence and three-dimensional structure. The combination of functional flexibility and the ability to make identical copies of each structure opens up many potential applications for crosslinked proteins as device components.

The authors demonstrated an application of their materials by using them to make dynamic enclosures for culturing cells (Fig. 1). The enclosures trap cells, expand as the cells multiply, and subsequently release the cultured cells in response to a change in pH. In the future, the properties of the protein gels used to make the microchambers could be modified by forming hybrid materials with synthetic polymers, and by using proteins that respond to a specific chemical signal. This approach has recently been demonstrated⁵ in a similar material, in which a modified version of the protein

calmodulin was trapped within a polymer network. The volume of the resulting gel changed when calcium ions or ligand molecules bound to the protein.

So what does the future hold for Kaehr and Shear's sophisticated protein gels? As potentially useful biomaterials, opportunities abound. The sequence, concentration and spatial presentation of biological signals within the material — such as binding pockets or the active sites of enzymes — can easily be varied, and the crosslinked proteins can be designed so that an external stimulus exposes such embedded features. And by using enzymes as a basis

for the gels, materials could be designed that amplify chemical signals or generate multiple responses from a single input.

Nevertheless, it won't all be plain sailing. Many proteins denature in response to changes in temperature or other environmental cues, and it isn't known whether all crosslinked proteins will undergo reversible changes in response to such signals. But as the repertoire of proteins that can be used in these materials expands, a tantalizing array of prospects can be envisaged: customized gels with passive and active molecular barriers to sift or release soluble chemical signals; systems that combine several design elements to mimic the complexity of naturally occurring tissues; or even actuators that have dynamic responses reminiscent of muscle contraction. ■

April M. Kloxin and Kristi S. Anseth are in the Department of Chemical and Biological Engineering, University of Colorado, Boulder, Colorado 80309-0424, USA.

e-mails: april.kloxin@colorado.edu; kristi.anseth@colorado.edu

1. Decker, C. J. *Coatings Technol.* **59**, 97–106 (1987).
2. Kaehr, B. & Shear, J. B. *Proc. Natl Acad. Sci. USA* **105**, 8850–8854 (2008).
3. Shen, H.-R., Spikes, J. D., Kopečeková, P. & Kopeček, J. *J. Photochem. Photobiol. B* **35**, 213–219 (1996).
4. Shen, H.-R., Spikes, J. D., Kopečeková, P. & Kopeček, J. *J. Photochem. Photobiol. B* **34**, 203–210 (1996).
5. Murphy, W. L., Dillmore, W. S., Modica, J. & Mrksich, M. *Angew. Chem. Int. Edn* **46**, 3066–3069 (2007).

ALZHEIMER'S DISEASE

The latest suspect

Rudolph E. Tanzi and Lars Bertram

Many genetic mutations and several environmental factors contribute to Alzheimer's disease. Yet another disease risk gene, one that is involved in calcium regulation, has been added to the mix.

Alzheimer's disease is characterized by neurodegeneration in several brain areas, with the hippocampus being one of the first regions affected. Accumulation of aggregates of the amyloid- β peptide, for example in the form of senile plaques, is another feature of this disorder. Dreses-Werringloer *et al.*¹ set out to search for genes that are preferentially expressed in the hippocampus and lie in chromosomal regions that have been linked² to the risk of developing Alzheimer's disease. Writing in *Cell*, they report the identification of a new Alzheimer's disease risk gene that not only meets both of these criteria, but also seems to modulate extracellular levels of amyloid- β .

The authors call this gene *CALHM1* (for calcium homeostasis modulator 1), because its product seems to be involved in calcium homeostasis. They show that increased expression of the *CALHM1* protein, which they find

to be localized in the cell membrane and in the membrane of an intracellular organelle known as the endoplasmic reticulum, enhances the entry of calcium into the cell and thus its concentration in the cytoplasm. They find that this results in changes in the processing of the amyloid precursor protein that lead to reduced levels of amyloid- β peptide.

Intrigued by these observations, the authors resequenced *CALHM1* from a small number of patients with Alzheimer's disease and healthy controls. They identify a single nucleotide polymorphism (variation) in this gene, which, at the protein level, results in the substitution of the proline amino acid at residue 86 with a leucine residue (P86L). In normal subjects, the gene version carrying this polymorphism (the leucine allele) seems to occur with a frequency of about 20%. In the patients with Alzheimer's, however, it occurs nearly twice as frequently. The authors then tested this polymorphism

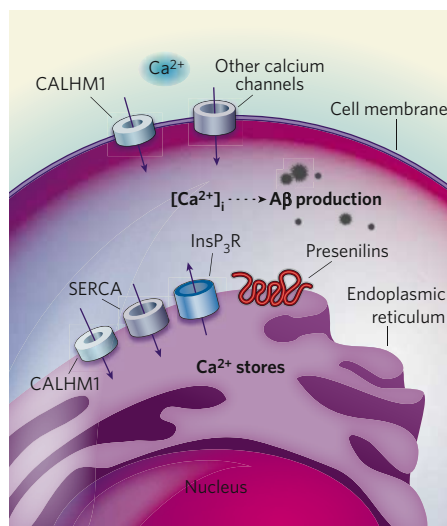


Figure 1 | Calcium homeostasis and Alzheimer's disease. Concentrations of calcium in the cytoplasm $[Ca^{2+}]_i$ are influenced by various types of calcium channel that span the cell membrane. They include store-operated capacitative-calcium-entry channels, receptor-operated channels and voltage-gated channels. Intracellular calcium levels are also coupled to calcium stores in the endoplasmic reticulum by presenilin proteins spanning the membrane of this organelle. Calcium stores in the endoplasmic reticulum are mainly regulated by the $InsP_3R$ receptor, the ryanodine receptor (not shown) and SERCA. Dreses-Werringloer *et al.*¹ identify a putative calcium channel, CALHM1, which occurs in both the cell membrane and the endoplasmic reticulum. They find that the leucine allele of the *CALHM1* gene, which is over-represented in samples from patients with Alzheimer's disease, leads to loss of CALHM1 function, decreases calcium permeability and reduces cytoplasmic calcium levels. Somehow, these events lead to increased production of amyloid- β peptide ($A\beta$).

in four other, independent groups of approximately 3,400 disease cases and controls and show that, in each data set, the leucine allele is consistently over-represented in Alzheimer's disease. Combining the data sets, they estimate that inheritance of the leucine allele increases the risk of developing Alzheimer's by about 40% — an effect roughly an order of magnitude smaller than the risk conferred by the $\epsilon 4$ allele of the only unequivocally established Alzheimer's disease risk gene *APOE*.

Dreses-Werringloer *et al.*¹ next show that the leucine allele leads to a loss of CALHM1 function, resulting in attenuated calcium permeability, reduced cytoplasmic calcium levels and increased amounts of extracellular amyloid- β . Several previous studies have also shown that intracellular calcium levels affect the metabolism of amyloid precursor protein. For example, inhibiting SERCA (a calcium channel in the endoplasmic-reticulum membrane) not only increases cytoplasmic levels of this ion, but also decreases amyloid- β levels³. Moreover, loss of function of the *PSEN* genes, which encode presenilin proteins (defects in these

proteins have been linked to familial forms of Alzheimer's disease), also affects cytoplasmic calcium levels. It activates capacitative calcium entry — a process that replenishes and regulates intracellular calcium stores in response to activation of particular cell-membrane calcium channels⁴. In turn, capacitative calcium entry leads to reduced levels of a form of amyloid- β called amyloid- $\beta 42$, which is more prone to plaque formation than the more common form, amyloid- $\beta 40$ (ref. 4). The same study also found that the mutations in the *PSEN1* and *PSEN2* genes that are linked^{5,6} to the familial form of the disease reduce capacitative calcium entry, leading to a selective increase in levels of amyloid- $\beta 42$.

Given that presenilins influence capacitative calcium entry, these molecules probably modulate the coupling between the calcium channels in the endoplasmic reticulum and those in the cell membrane. Dreses-Werringloer and colleagues' observations¹ indicate that the Alzheimer's disease-associated P86L polymorphism in *CALHM1* might also disrupt coupling between these two calcium regulators, leading to increased amyloid- β levels (Fig. 1). It is unclear exactly how intracellular calcium levels regulate the processing of amyloid precursor protein and amyloid- β production.

Another new study⁷ also investigates modulations in calcium homeostasis in Alzheimer's disease. These authors find that the same mutations in *PSEN1* and *PSEN2* as those identified for their effects on capacitative calcium entry⁴ stimulate the gating activity of a calcium-channel receptor called $InsP_3R$, and enhance its sensitivity to the signalling molecule $InsP_3$. In agreement with earlier observations⁴, the effects of these mutations on calcium homeostasis seem to be linked to elevated amyloid- $\beta 42$ levels. Collectively, these findings^{1,4,7} indicate a link between the risk of developing Alzheimer's disease and mutations/polymorphisms in genes that can modulate cellular calcium stores (for a review see ref. 8).

Despite Dreses-Werringloer and colleagues' observations, it is too early to declare *CALHM1* an 'Alzheimer's disease gene'. Looking back at research into the genetics of this disease, hundreds of papers have reported data for dozens of 'novel Alzheimer's disease genes' over the past three decades. But the vast majority of such studies did not stand the test of time, with independent work failing to replicate them². Chromosomal region 10q11–q25 has been the subject of several converging genetic linkage analyses⁹ (which assess the link between specific genetic markers in a chromosomal region and the risk of a disease). Such analyses have made this region the focus of intensive genetic investigation in relation to Alzheimer's disease. In this region alone, 80 genes and loci have been assessed for their association with this disorder in nearly 250 publications. Among them are candidates (such as *CHAT*, *CH25H*, *IDE*, *PLAU*, *SORCS1* and *VR22*) whose claim to a potential functional involvement in Alzheimer's



50 YEARS AGO

This year at the European Centre for Nuclear Research (C.E.R.N.) the [annual international] conference was divided into ten half-day sessions and all contributions in a particular field were presented to the conference by one person, the rapporteur, whose job it was to introduce the subject and to present and to explain the significance of the results obtained ... The success of this procedure depends very much on the rapporteur ... The general opinion at the end of the conference was that the new method of presentation had been enough of a success for the method to be tried again. From *Nature* 9 August 1958.

100 YEARS AGO

Refrigeration: an Elementary Text-book. By J. Wemyss Anderson — The increasing use of refrigerating processes in the distribution and preservation of food, and also in many important industries, has already called for a special type of engineer who must possess a knowledge, not only of machines and mechanism, but also of the theoretical properties of heat. Nowadays, when the market for electrical engineers is becoming uncomfortably crowded, young men would do well to consider the prospects open to them as refrigerating engineers ... A very important branch of the subject is dealt with in chapter x, where the methods of insulating and cooling large chambers are described and illustrated.

ALSO:

To the July number of the *Zoologist* Mr. S. M. Perlmann contributes a paper entitled "Is the Okapi Identical with the 'Thahash' of the Jews?" The word "thahash" has been translated "badger" and "dolphin", but both these readings are conjectural. The Talmudists, according to Mr. Perlmann, considered it, however, to be a ruminant, of a beautiful colour, with a single horn on the forehead.

From *Nature* 6 August 1908.

50 & 100 YEARS AGO

disease is no less compelling than that of *CALHM1*. Systematic meta-analyses, performed as part of the AlzGene database project^{2,10}, show nominally significant association with Alzheimer's disease for a number of these loci. But these genes are still far from representing 'established' Alzheimer's disease genes, a title currently bestowed only on *APOE*. At present, they are no more than the best bets across a vast body of genetic data, with *CALHM1* representing the latest addition.

Only further independent replication studies will establish whether *CALHM1* is a bona fide Alzheimer's disease gene. In that regard, it is concerning that two genome-wide association studies of the disease^{11,12} (based on a reanalysis of publicly available genotype data) reveal no evidence of an association between the disease

and single nucleotide polymorphisms in the chromosomal region containing *CALHM1* in 2,900 subjects — disease cases and controls combined.

Moreover, the data sets from these studies^{11,12} contain two single nucleotide polymorphisms whose less-frequent alleles are consistently co-inherited with the leucine allele in *CALHM1*. These polymorphisms can therefore be considered reasonable proxies for the P86L polymorphism. In contrast with Dreeses-Werringloer and colleagues' observation¹, however, neither of these polymorphisms is significantly associated with the risk of Alzheimer's disease. Such early negative-association results suggest that *CALHM1*'s road to becoming an established Alzheimer's disease gene may be as rutted as that encountered by many previous candidates. ■

Rudolph E. Tanzi and Lars Bertram are in the Genetics and Aging Research Unit, MassGeneral Institute for Neurodegenerative Disease (MIND), Department of Neurology, Massachusetts General Hospital, Charlestown, Massachusetts 02129, USA.

e-mail: tanzi@helix.mgh.harvard.edu

1. Dreeses-Werringloer, U. *et al.* *Cell* **133**, 1149–1161 (2008).
2. Bertram, L., McQueen, M. B., Mullin, K., Blacker, D. & Tanzi, R. E. *Nature Genet.* **39**, 17–23 (2007).
3. Green, K. N. *et al.* *J. Cell Biol.* **181**, 1107–1116 (2008).
4. Yoo, A. S. *et al.* *Neuron* **27**, 561–572 (2000).
5. Sherrington, R. *et al.* *Nature* **375**, 754–760 (1995).
6. Levy-Lahad, E. *et al.* *Science* **269**, 973–977 (1995).
7. Cheung, K.-H. *et al.* *Neuron* **58**, 871–883 (2008).
8. LaFerla, F. M. *Nature Rev. Neurosci.* **3**, 862–872 (2002).
9. Bertram, L. *et al.* *Science* **290**, 2302–2303 (2000).
10. www.alzgene.org
11. Reiman, E. M. *et al.* *Neuron* **54**, 713–720 (2007).
12. Li, H. *et al.* *Arch. Neurol.* **65**, 45–53 (2008).

NANOTECHNOLOGY

Diamonds are for tethers

Robert J. Hamers

Modified diamond nanowires produce an electrical response on binding to DNA. This gem of a discovery could pave the way to robust biosensors that use electrical signals to detect molecules.

Diamonds may well be a girl's best friend. But some people prefer to sing the praises of their extreme chemical stability, mechanical robustness and tunable electronic properties. The last of these characteristics has led to impressive advances in the use of diamonds for sensing applications. Reporting in *Angewandte Chemie*, Yang *et al.*¹ describe a method for making biosensors from diamond nanowires. Their devices produce an electrical signal on binding to DNA molecules, and are so sensitive that they can detect vanishingly small amounts of the target molecules (picomolar concentrations).

Diamond has remarkable properties that make it potentially useful for many applications, but for many years its high cost prohibited its exploitation. Fortunately, the development of methods for synthesizing diamond, both in bulk and as thin films, has made the material much cheaper, leading to an explosion in research and several notable breakthroughs. Most relevant to Yang and colleagues' work¹ was the report that monolayers of organic molecules covalently grafted to a diamond surface have been used to bind short strands of DNA, with excellent biological selectivity^{2,3}. Subsequent studies^{4–8} have shown that diamond surfaces decorated

with DNA or proteins can form the basis of real-time biosensors, using several different approaches.

The common theme of these diamond biosensors is that they all convert biological information — the presence or absence of DNA binding to the surface — into an electrical signal.

Such signals can be measured accurately, and the devices that measure them can easily be scaled up to incorporate large numbers of sensor components^{9,10}. Furthermore, the technology required to create a biosensor based on electrical signals is much cheaper than that required for fluorescence-based sensors, which have so far dominated the field.

Biological-to-electrical signal conversion is attractive for sensors that are in continuous use or that need to withstand harsh conditions, as would be the case for real-time environmental monitoring. This is because such sensors would be far more robust than their fluorescence-based counterparts, which require the use of fragile fluorescently labelled molecules. But so far, sensors that use electrical signals have typically been thousands of times less sensitive than the best fluorescence-based detection systems.

Yang *et al.*¹ now demonstrate that the detection limits of electrical biosensors are markedly improved if highly conductive diamond nanowires, vertically aligned on a diamond surface like a minuscule bed of nails (Fig. 1), are used as substrates for DNA-detection elements. To make the nanowires, the authors modified a procedure¹¹ known as reactive ion etching, in which parts of a diamond surface are selectively removed to form three-dimensional structures. They first deposited diamond nanoparticles on a diamond surface, to protect the underlying regions. Subsequent etching left behind pointed structures in the diamond surface — nanowires, or perhaps more accurately, nanocones. The wires were about 10 nanometres high, and were spaced about the same distance apart.

Next, the authors tethered short, single-stranded DNA molecules to the nanowires. They then used

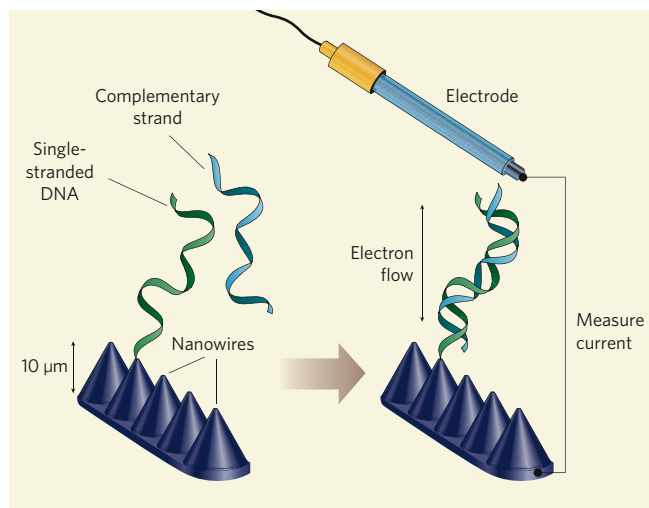


Figure 1 | DNA sensing on diamond nanowires. Yang *et al.*¹ have prepared DNA sensors by attaching single strands of DNA to diamond nanowires that have been constructed on a diamond surface. Both DNA and diamond conduct electricity, so electrons from the diamond substrate can flow along the DNA. The conductivity of the system changes when complementary strands of DNA bind to the tethered DNA. This effect can be quantified by immersing the diamond sensor in a solution of DNA, placing an electrode close to the diamond surface and measuring the current that flows between them.

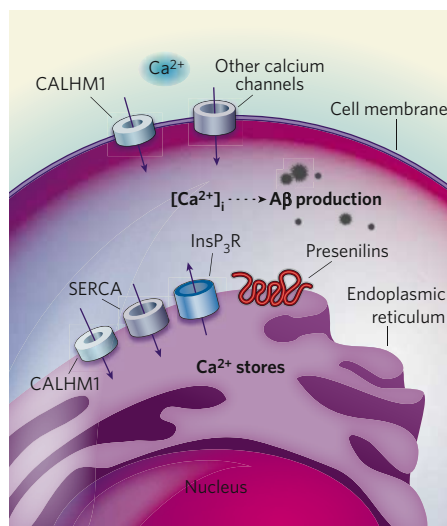


Figure 1 | Calcium homeostasis and Alzheimer's disease. Concentrations of calcium in the cytoplasm $[Ca^{2+}]_i$ are influenced by various types of calcium channel that span the cell membrane. They include store-operated capacitative-calcium-entry channels, receptor-operated channels and voltage-gated channels. Intracellular calcium levels are also coupled to calcium stores in the endoplasmic reticulum by presenilin proteins spanning the membrane of this organelle. Calcium stores in the endoplasmic reticulum are mainly regulated by the $InsP_3R$ receptor, the ryanodine receptor (not shown) and SERCA. Dreses-Werringloer *et al.*¹ identify a putative calcium channel, CALHM1, which occurs in both the cell membrane and the endoplasmic reticulum. They find that the leucine allele of the *CALHM1* gene, which is over-represented in samples from patients with Alzheimer's disease, leads to loss of CALHM1 function, decreases calcium permeability and reduces cytoplasmic calcium levels. Somehow, these events lead to increased production of amyloid- β peptide ($A\beta$).

in four other, independent groups of approximately 3,400 disease cases and controls and show that, in each data set, the leucine allele is consistently over-represented in Alzheimer's disease. Combining the data sets, they estimate that inheritance of the leucine allele increases the risk of developing Alzheimer's by about 40% — an effect roughly an order of magnitude smaller than the risk conferred by the $\epsilon 4$ allele of the only unequivocally established Alzheimer's disease risk gene *APOE*.

Dreses-Werringloer *et al.*¹ next show that the leucine allele leads to a loss of CALHM1 function, resulting in attenuated calcium permeability, reduced cytoplasmic calcium levels and increased amounts of extracellular amyloid- β . Several previous studies have also shown that intracellular calcium levels affect the metabolism of amyloid precursor protein. For example, inhibiting SERCA (a calcium channel in the endoplasmic-reticulum membrane) not only increases cytoplasmic levels of this ion, but also decreases amyloid- β levels³. Moreover, loss of function of the *PSEN* genes, which encode presenilin proteins (defects in these

proteins have been linked to familial forms of Alzheimer's disease), also affects cytoplasmic calcium levels. It activates capacitative calcium entry — a process that replenishes and regulates intracellular calcium stores in response to activation of particular cell-membrane calcium channels⁴. In turn, capacitative calcium entry leads to reduced levels of a form of amyloid- β called amyloid- $\beta 42$, which is more prone to plaque formation than the more common form, amyloid- $\beta 40$ (ref. 4). The same study also found that the mutations in the *PSEN1* and *PSEN2* genes that are linked^{5,6} to the familial form of the disease reduce capacitative calcium entry, leading to a selective increase in levels of amyloid- $\beta 42$.

Given that presenilins influence capacitative calcium entry, these molecules probably modulate the coupling between the calcium channels in the endoplasmic reticulum and those in the cell membrane. Dreses-Werringloer and colleagues' observations¹ indicate that the Alzheimer's disease-associated P86L polymorphism in *CALHM1* might also disrupt coupling between these two calcium regulators, leading to increased amyloid- β levels (Fig. 1). It is unclear exactly how intracellular calcium levels regulate the processing of amyloid precursor protein and amyloid- β production.

Another new study⁷ also investigates modulations in calcium homeostasis in Alzheimer's disease. These authors find that the same mutations in *PSEN1* and *PSEN2* as those identified for their effects on capacitative calcium entry⁴ stimulate the gating activity of a calcium-channel receptor called $InsP_3R$, and enhance its sensitivity to the signalling molecule $InsP_3$. In agreement with earlier observations⁴, the effects of these mutations on calcium homeostasis seem to be linked to elevated amyloid- $\beta 42$ levels. Collectively, these findings^{1,4,7} indicate a link between the risk of developing Alzheimer's disease and mutations/polymorphisms in genes that can modulate cellular calcium stores (for a review see ref. 8).

Despite Dreses-Werringloer and colleagues' observations, it is too early to declare *CALHM1* an 'Alzheimer's disease gene'. Looking back at research into the genetics of this disease, hundreds of papers have reported data for dozens of 'novel Alzheimer's disease genes' over the past three decades. But the vast majority of such studies did not stand the test of time, with independent work failing to replicate them². Chromosomal region 10q11–q25 has been the subject of several converging genetic linkage analyses⁹ (which assess the link between specific genetic markers in a chromosomal region and the risk of a disease). Such analyses have made this region the focus of intensive genetic investigation in relation to Alzheimer's disease. In this region alone, 80 genes and loci have been assessed for their association with this disorder in nearly 250 publications. Among them are candidates (such as *CHAT*, *CH25H*, *IDE*, *PLAU*, *SORCS1* and *VR22*) whose claim to a potential functional involvement in Alzheimer's



50 YEARS AGO

This year at the European Centre for Nuclear Research (C.E.R.N.) the [annual international] conference was divided into ten half-day sessions and all contributions in a particular field were presented to the conference by one person, the rapporteur, whose job it was to introduce the subject and to present and to explain the significance of the results obtained ... The success of this procedure depends very much on the rapporteur ... The general opinion at the end of the conference was that the new method of presentation had been enough of a success for the method to be tried again. From *Nature* 9 August 1958.

100 YEARS AGO

Refrigeration: an Elementary Text-book. By J. Wemyss Anderson — The increasing use of refrigerating processes in the distribution and preservation of food, and also in many important industries, has already called for a special type of engineer who must possess a knowledge, not only of machines and mechanism, but also of the theoretical properties of heat. Nowadays, when the market for electrical engineers is becoming uncomfortably crowded, young men would do well to consider the prospects open to them as refrigerating engineers ... A very important branch of the subject is dealt with in chapter x, where the methods of insulating and cooling large chambers are described and illustrated.

ALSO:

To the July number of the *Zoologist* Mr. S. M. Perlmann contributes a paper entitled "Is the Okapi Identical with the 'Thahash' of the Jews?" The word "thahash" has been translated "badger" and "dolphin", but both these readings are conjectural. The Talmudists, according to Mr. Perlmann, considered it, however, to be a ruminant, of a beautiful colour, with a single horn on the forehead.

From *Nature* 6 August 1908.

50 & 100 YEARS AGO

disease is no less compelling than that of *CALHM1*. Systematic meta-analyses, performed as part of the AlzGene database project^{2,10}, show nominally significant association with Alzheimer's disease for a number of these loci. But these genes are still far from representing 'established' Alzheimer's disease genes, a title currently bestowed only on *APOE*. At present, they are no more than the best bets across a vast body of genetic data, with *CALHM1* representing the latest addition.

Only further independent replication studies will establish whether *CALHM1* is a bona fide Alzheimer's disease gene. In that regard, it is concerning that two genome-wide association studies of the disease^{11,12} (based on a reanalysis of publicly available genotype data) reveal no evidence of an association between the disease

and single nucleotide polymorphisms in the chromosomal region containing *CALHM1* in 2,900 subjects — disease cases and controls combined.

Moreover, the data sets from these studies^{11,12} contain two single nucleotide polymorphisms whose less-frequent alleles are consistently co-inherited with the leucine allele in *CALHM1*. These polymorphisms can therefore be considered reasonable proxies for the P86L polymorphism. In contrast with Dreeses-Werringloer and colleagues' observation¹, however, neither of these polymorphisms is significantly associated with the risk of Alzheimer's disease. Such early negative-association results suggest that *CALHM1*'s road to becoming an established Alzheimer's disease gene may be as rutted as that encountered by many previous candidates. ■

Rudolph E. Tanzi and Lars Bertram are in the Genetics and Aging Research Unit, MassGeneral Institute for Neurodegenerative Disease (MIND), Department of Neurology, Massachusetts General Hospital, Charlestown, Massachusetts 02129, USA.

e-mail: tanzi@helix.mgh.harvard.edu

1. Dreeses-Werringloer, U. *et al. Cell* **133**, 1149–1161 (2008).
2. Bertram, L., McQueen, M. B., Mullin, K., Blacker, D. & Tanzi, R. E. *Nature Genet.* **39**, 17–23 (2007).
3. Green, K. N. *et al. J. Cell Biol.* **181**, 1107–1116 (2008).
4. Yoo, A. S. *et al. Neuron* **27**, 561–572 (2000).
5. Sherrington, R. *et al. Nature* **375**, 754–760 (1995).
6. Levy-Lahad, E. *et al. Science* **269**, 973–977 (1995).
7. Cheung, K.-H. *et al. Neuron* **58**, 871–883 (2008).
8. LaFerla, F. M. *Nature Rev. Neurosci.* **3**, 862–872 (2002).
9. Bertram, L. *et al. Science* **290**, 2302–2303 (2000).
10. www.alzgene.org
11. Reiman, E. M. *et al. Neuron* **54**, 713–720 (2007).
12. Li, H. *et al. Arch. Neurol.* **65**, 45–53 (2008).

NANOTECHNOLOGY

Diamonds are for tethers

Robert J. Hamers

Modified diamond nanowires produce an electrical response on binding to DNA. This gem of a discovery could pave the way to robust biosensors that use electrical signals to detect molecules.

Diamonds may well be a girl's best friend. But some people prefer to sing the praises of their extreme chemical stability, mechanical robustness and tunable electronic properties. The last of these characteristics has led to impressive advances in the use of diamonds for sensing applications. Reporting in *Angewandte Chemie*, Yang *et al.*¹ describe a method for making biosensors from diamond nanowires. Their devices produce an electrical signal on binding to DNA molecules, and are so sensitive that they can detect vanishingly small amounts of the target molecules (picomolar concentrations).

Diamond has remarkable properties that make it potentially useful for many applications, but for many years its high cost prohibited its exploitation. Fortunately, the development of methods for synthesizing diamond, both in bulk and as thin films, has made the material much cheaper, leading to an explosion in research and several notable breakthroughs. Most relevant to Yang and colleagues' work¹ was the report that monolayers of organic molecules covalently grafted to a diamond surface have been used to bind short strands of DNA, with excellent biological selectivity^{2,3}. Subsequent studies^{4–8} have shown that diamond surfaces decorated

with DNA or proteins can form the basis of real-time biosensors, using several different approaches.

The common theme of these diamond biosensors is that they all convert biological information — the presence or absence of DNA binding to the surface — into an electrical signal.

Such signals can be measured accurately, and the devices that measure them can easily be scaled up to incorporate large numbers of sensor components^{9,10}. Furthermore, the technology required to create a biosensor based on electrical signals is much cheaper than that required for fluorescence-based sensors, which have so far dominated the field.

Biological-to-electrical signal conversion is attractive for sensors that are in continuous use or that need to withstand harsh conditions, as would be the case for real-time environmental monitoring. This is because such sensors would be far more robust than their fluorescence-based counterparts, which require the use of fragile fluorescently labelled molecules. But so far, sensors that use electrical signals have typically been thousands of times less sensitive than the best fluorescence-based detection systems.

Yang *et al.*¹ now demonstrate that the detection limits of electrical biosensors are markedly improved if highly conductive diamond nanowires, vertically aligned on a diamond surface like a minuscule bed of nails (Fig. 1), are used as substrates for DNA-detection elements. To make the nanowires, the authors modified a procedure¹¹ known as reactive ion etching, in which parts of a diamond surface are selectively removed to form three-dimensional structures. They first deposited diamond nanoparticles on a diamond surface, to protect the underlying regions. Subsequent etching left behind pointed structures in the diamond surface — nanowires, or perhaps more accurately, nanocones. The wires were about 10 nanometres high, and were spaced about the same distance apart.

Next, the authors tethered short, single-stranded DNA molecules to the nanowires. They then used

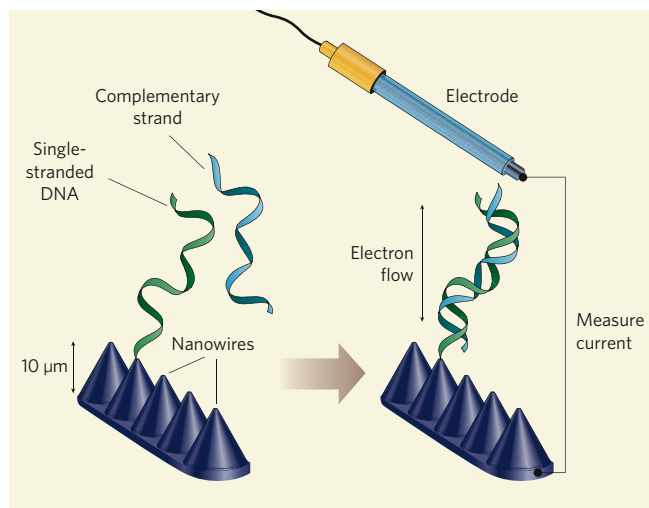


Figure 1 | DNA sensing on diamond nanowires. Yang *et al.*¹ have prepared DNA sensors by attaching single strands of DNA to diamond nanowires that have been constructed on a diamond surface. Both DNA and diamond conduct electricity, so electrons from the diamond substrate can flow along the DNA. The conductivity of the system changes when complementary strands of DNA bind to the tethered DNA. This effect can be quantified by immersing the diamond sensor in a solution of DNA, placing an electrode close to the diamond surface and measuring the current that flows between them.

electrochemical measurements to monitor changes in the electronic properties of the diamond surface when single-stranded DNA molecules, complementary in structure to that of the tethered DNA, were trapped from solution by the nanowire-supported DNA. Their results showed that trapping complementary DNA interferes with electron-transfer processes in the nanowires, thereby providing an electrical read-out. The resulting sensor could detect exceptionally low levels of the DNA analyte.

So how is the electrical signal in Yang and colleagues' method generated? Several mechanisms could apply. Diamond is a semiconductor, which means that its electrical properties can be altered by externally applied electric fields. DNA molecules are negatively charged, and so, when bound close to a diamond surface, they can induce an electric field that alters the conductivity of the diamond in what is known as a field effect. Another possibility depends on the conductivity of DNA itself. Hybridization of single-stranded DNA into the double-stranded form leads to an increase in the conductivity of the molecules, an effect that is generally attributed to the closer proximity of the electron systems of the DNA bases in the duplex¹².

Yang *et al.*¹ used several electrochemical measurements to try to understand the physical origin of the electrical signals, with conflicting results. Their cyclic voltammetry experiments (which measured the current through their sensor in response to voltage) showed that DNA hybridization decreased the conductivity of the nanowires, whereas impedance measurements (which characterized the electrical response of the system at specific frequencies of alternating current) suggested the reverse effect. The reason for this discrepancy is unknown, but will probably be a consequence of the electron-transfer and diffusion processes that occur at the surfaces of complex nanostructured materials. More work is clearly needed to understand this. Unravelling the origins of the electrical signals will also be crucial to developing robust analytical devices in the future.

Ultimately, the sensitivity of biosensors that rely on surface-derivatized components is limited by the physical parameters that govern the adsorption of DNA to those surfaces¹³. Similar limits of detection to that of Yang and colleagues' biosensor¹ have been achieved by diamond-based field-effect transistors⁶, suggesting that there may be several approaches to converting biological signals into electrical ones that take advantage of diamond's extraordinary properties. Nevertheless, the authors' discovery might trigger the development of sensors for clinical diagnostics, environmental sensing and other applications at the interface between biology and microelectronics. ■

Robert J. Hamers is in the Department of Chemistry, University of Wisconsin–Madison, 1101 University Avenue, Madison, Wisconsin 53706, USA.

e-mail: rjhamers@wisc.edu

1. Yang, N., Uetsuka, H., Osawa, E. & Nebel, C. E. *Angew. Chem. Int. Edn* **47**, 5183–5185 (2008).
2. Yang, W. *et al. Nature Mater.* **1**, 253–257 (2002).
3. Yang, W. *et al. Chem. Mater.* **17**, 938–940 (2005).
4. Härtl, A. *et al. Nature Mater.* **3**, 736–742 (2004).
5. Yang, W., Butler, J. E., Russell, J. N. Jr & Hamers, R. J. *Analyst* **132**, 296–306 (2007).
6. Song, K.-S. *et al. Phys. Rev. E* **74**, 041919 (2006).
7. Song, K.-S., Hiraki, T., Umezawa, H. & Kawarada, H. *Appl. Phys. Lett.* **90**, 063901 (2007).

8. Yang, W. & Hamers, R. J. *Appl. Phys. Lett.* **85**, 3626–3628 (2004).
9. Katz, E. & Willner, I. *Electroanalysis* **15**, 913–947 (2003).
10. Wang, J. *Chem. Eur. J.* **5**, 1681–1685 (1999).
11. Zou, Y. S. *et al. Appl. Phys. Lett.* **92**, 053105 (2008).
12. Boon, E. M., Ceres, D. M., Drummond, T. G., Hill, M. G. & Barton, J. K. *Nature Biotechnol.* **18**, 1096–1100 (2000).
13. Lee, H. J., Li, Y., Wark, A. W. & Corn, R. M. *Anal. Chem.* **77**, 5096–5100 (2005).

CELL BIOLOGY

A molecular age barrier

Matt Kaeberlein

A mother's instinct is to protect her children at any cost. In the budding yeast *Saccharomyces cerevisiae* this 'maternal instinct' comes at a high price — accelerated ageing and premature death.

Cells of budding yeast divide asymmetrically, with the larger mother cell easily distinguishable from her daughter. This asymmetry, which is not just structural but also affects the distribution of cellular components, ensures a type of ageing in yeast — replicative ageing — that is defined by the number of daughter cells a mother produces¹. Some years ago, a diffusible senescence factor associated with replicative ageing was discovered^{2,3}. Over most of the course of the mother cell's lifespan, this factor

is successfully retained in the mother, allowing each daughter cell to begin life relatively free of age-associated damage. In very old mother cells, however, the 'age barrier' becomes overburdened, leading to loss of asymmetry and premature daughter-cell senescence. On page 728 of this issue, Shcheprova *et al.*⁴ provide insight into a nuclear age barrier in yeast, and show that, surprisingly, it limits longevity in both mother and daughter cells.

Several types of molecular damage are

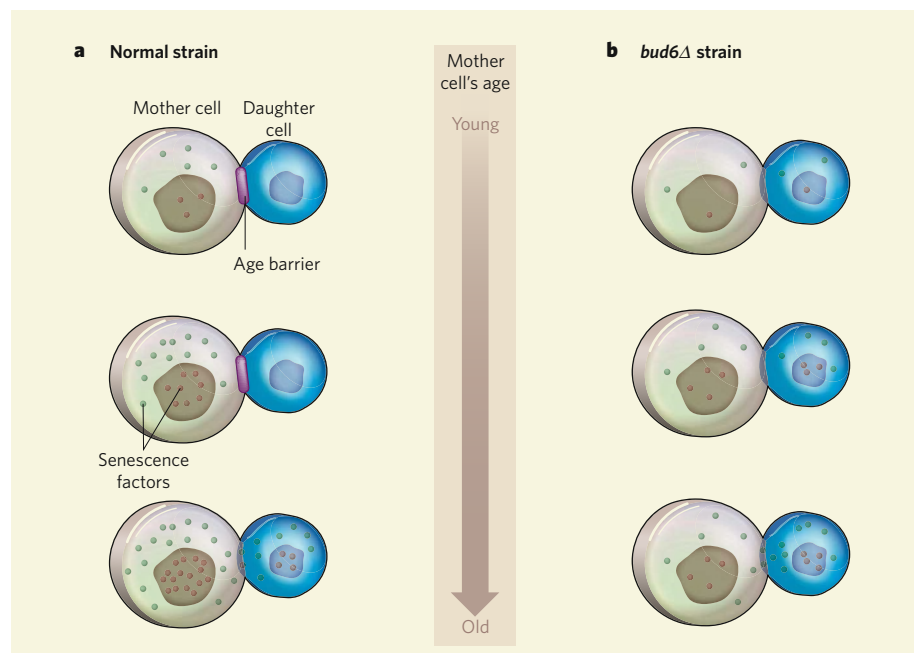


Figure 1 | Shared ageing. During cell division, normal young yeast mother cells asymmetrically retain a diffusible senescence factor. This factor may be a composite of damage-inducing nuclear factors, including extrachromosomal ribosomal-DNA circles (ERCs), and cytoplasmic components such as oxidatively damaged proteins and dysfunctional mitochondria. On the basis of Shcheprova and colleagues' identification⁴ of a nuclear age barrier in yeast, which limits the passage of ERCs from mother cell to daughter cell, a general model can be proposed. **a**, During replicative ageing of normal yeast cells, an age barrier prevents the passage of various senescence factors to daughter cells. But if the mother cell replicates very late in life, this barrier breaks down. **b**, In a mutant that lacks the age barrier (such as *bud6Δ*), population lifespan could be increased owing to equal distribution — and so decreased accumulation — of damage to both mother and daughter cells.

electrochemical measurements to monitor changes in the electronic properties of the diamond surface when single-stranded DNA molecules, complementary in structure to that of the tethered DNA, were trapped from solution by the nanowire-supported DNA. Their results showed that trapping complementary DNA interferes with electron-transfer processes in the nanowires, thereby providing an electrical read-out. The resulting sensor could detect exceptionally low levels of the DNA analyte.

So how is the electrical signal in Yang and colleagues' method generated? Several mechanisms could apply. Diamond is a semiconductor, which means that its electrical properties can be altered by externally applied electric fields. DNA molecules are negatively charged, and so, when bound close to a diamond surface, they can induce an electric field that alters the conductivity of the diamond in what is known as a field effect. Another possibility depends on the conductivity of DNA itself. Hybridization of single-stranded DNA into the double-stranded form leads to an increase in the conductivity of the molecules, an effect that is generally attributed to the closer proximity of the electron systems of the DNA bases in the duplex¹².

Yang *et al.*¹ used several electrochemical measurements to try to understand the physical origin of the electrical signals, with conflicting results. Their cyclic voltammetry experiments (which measured the current through their sensor in response to voltage) showed that DNA hybridization decreased the conductivity of the nanowires, whereas impedance measurements (which characterized the electrical response of the system at specific frequencies of alternating current) suggested the reverse effect. The reason for this discrepancy is unknown, but will probably be a consequence of the electron-transfer and diffusion processes that occur at the surfaces of complex nanostructured materials. More work is clearly needed to understand this. Unravelling the origins of the electrical signals will also be crucial to developing robust analytical devices in the future.

Ultimately, the sensitivity of biosensors that rely on surface-derivatized components is limited by the physical parameters that govern the adsorption of DNA to those surfaces¹³. Similar limits of detection to that of Yang and colleagues' biosensor¹ have been achieved by diamond-based field-effect transistors⁶, suggesting that there may be several approaches to converting biological signals into electrical ones that take advantage of diamond's extraordinary properties. Nevertheless, the authors' discovery might trigger the development of sensors for clinical diagnostics, environmental sensing and other applications at the interface between biology and microelectronics. ■

Robert J. Hamers is in the Department of Chemistry, University of Wisconsin–Madison, 1101 University Avenue, Madison, Wisconsin 53706, USA.

e-mail: rjhamers@wisc.edu

1. Yang, N., Uetsuka, H., Osawa, E. & Nebel, C. E. *Angew. Chem. Int. Edn* **47**, 5183–5185 (2008).
2. Yang, W. *et al. Nature Mater.* **1**, 253–257 (2002).
3. Yang, W. *et al. Chem. Mater.* **17**, 938–940 (2005).
4. Härtl, A. *et al. Nature Mater.* **3**, 736–742 (2004).
5. Yang, W., Butler, J. E., Russell, J. N. Jr & Hamers, R. J. *Analyst* **132**, 296–306 (2007).
6. Song, K.-S. *et al. Phys. Rev. E* **74**, 041919 (2006).
7. Song, K.-S., Hiraki, T., Umezawa, H. & Kawarada, H. *Appl. Phys. Lett.* **90**, 063901 (2007).

8. Yang, W. & Hamers, R. J. *Appl. Phys. Lett.* **85**, 3626–3628 (2004).
9. Katz, E. & Willner, I. *Electroanalysis* **15**, 913–947 (2003).
10. Wang, J. *Chem. Eur. J.* **5**, 1681–1685 (1999).
11. Zou, Y. S. *et al. Appl. Phys. Lett.* **92**, 053105 (2008).
12. Boon, E. M., Ceres, D. M., Drummond, T. G., Hill, M. G. & Barton, J. K. *Nature Biotechnol.* **18**, 1096–1100 (2000).
13. Lee, H. J., Li, Y., Wark, A. W. & Corn, R. M. *Anal. Chem.* **77**, 5096–5100 (2005).

CELL BIOLOGY

A molecular age barrier

Matt Kaeberlein

A mother's instinct is to protect her children at any cost. In the budding yeast *Saccharomyces cerevisiae* this 'maternal instinct' comes at a high price — accelerated ageing and premature death.

Cells of budding yeast divide asymmetrically, with the larger mother cell easily distinguishable from her daughter. This asymmetry, which is not just structural but also affects the distribution of cellular components, ensures a type of ageing in yeast — replicative ageing — that is defined by the number of daughter cells a mother produces¹. Some years ago, a diffusible senescence factor associated with replicative ageing was discovered^{2,3}. Over most of the course of the mother cell's lifespan, this factor

is successfully retained in the mother, allowing each daughter cell to begin life relatively free of age-associated damage. In very old mother cells, however, the 'age barrier' becomes overburdened, leading to loss of asymmetry and premature daughter-cell senescence. On page 728 of this issue, Shcheprova *et al.*⁴ provide insight into a nuclear age barrier in yeast, and show that, surprisingly, it limits longevity in both mother and daughter cells.

Several types of molecular damage are

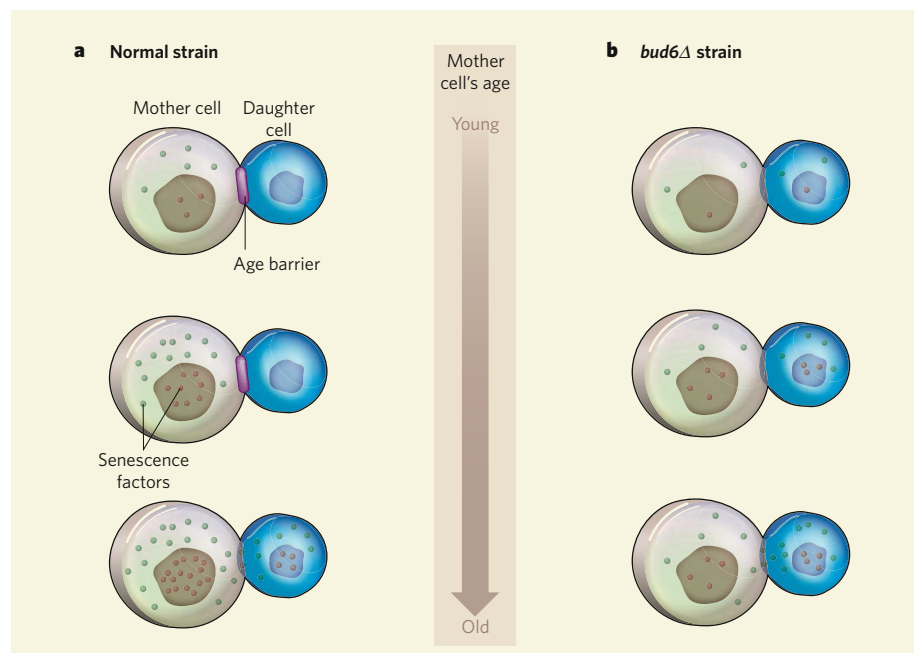


Figure 1 | Shared ageing. During cell division, normal young yeast mother cells asymmetrically retain a diffusible senescence factor. This factor may be a composite of damage-inducing nuclear factors, including extrachromosomal ribosomal-DNA circles (ERCs), and cytoplasmic components such as oxidatively damaged proteins and dysfunctional mitochondria. On the basis of Shcheprova and colleagues' identification⁴ of a nuclear age barrier in yeast, which limits the passage of ERCs from mother cell to daughter cell, a general model can be proposed. **a**, During replicative ageing of normal yeast cells, an age barrier prevents the passage of various senescence factors to daughter cells. But if the mother cell replicates very late in life, this barrier breaks down. **b**, In a mutant that lacks the age barrier (such as *bud6Δ*), population lifespan could be increased owing to equal distribution — and so decreased accumulation — of damage to both mother and daughter cells.

thought to contribute to ageing in yeast. The best studied of these are extrachromosomal ribosomal-DNA circles (ERCs) — self-replicating plasmids formed by the process of homologous recombination within ribosomal DNA⁵. Consistent with their role as a senescence factor, ERCs are preferentially retained in the mother-cell nucleus and accumulate with age⁵. Mutations that accelerate the rate of ERC formation shorten replicative lifespan, whereas those that reduce ERC accumulation enhance longevity.

Besides ERCs, dysfunctional mitochondria and proteins damaged by reactive oxygen species are actively retained by mother cells and have also been implicated in replicative ageing^{6–8}. Although the importance of such cytoplasmic factors in yeast ageing is less established than that of ERCs, it seems that the accumulation of damage in several mother-cell compartments results in replicative ageing.

Shcheprova *et al.*⁴ advance our understanding of asymmetry between yeast mother and daughter cells, and its importance in ageing, by identifying a diffusion barrier that compartmentalizes the nuclear envelope during budding. This barrier's formation depends on septin proteins, which assemble into a ring around the bud neck and prevent the passage of pre-existing nuclear pores to the daughter cell. Circular DNA molecules (such as ERCs) — and presumably other asymmetrically inherited nuclear factors — are associated with these nuclear-pore complexes, and so are also unable to cross the barrier. This partitioning process requires Bud6, a protein that localizes to the bud neck. Yeast mother cells — and daughter cells — in which the gene encoding this protein has been deleted (*bud6Δ* cells) no longer retain pre-existing nuclear-pore complexes and ERCs, and are consequently long-lived (Fig. 1). This indicates that the presence of the nuclear age barrier limits longevity.

The authors' observations raise some perplexing and intriguing questions. For example, it is not clear why loss of ERCs' asymmetrical distribution increases replicative lifespan. If *bud6Δ* daughter cells inherit ERCs at birth, it seems counterintuitive that they should be long-lived. This apparent paradox can be explained if the ability of an ageing mother cell to pass ERCs to daughter cells more than compensates for the ERCs she herself inherited at birth (Fig. 1). Accordingly, although the daughter cells that arise from young *bud6Δ* mother cells (those that have been through fewer than about eight replicative cycles) inherit some damage at birth, they also go on to continually transfer damage to each of their own daughter cells and, consequently, live longer than do normal cells. Because ERCs and possibly other ageing factors accumulate with age, daughter cells arising from aged *bud6Δ* mother cells (those that have been through more than eight replicative cycles) will inherit a damage burden sufficient to shorten their lifespan. But as the proportion of mother cells of this age or older in

PALAEONTOLOGY

Bite size

Ten thousand or so years ago, the lineage of sabre-toothed cats came to an end. Per Christiansen has revisited the evolutionary history of this distinctive group and that of the other members (felines) of the cat family, with particular reference to changes in the shape of their skull and mandible. He describes his work, which bears on the lively issue of the cause of the sabretooths' demise, in *PLoS ONE* (P. Christiansen *PLoS ONE* **3**, e2807; 2008). It scarcely needs pointing out that the characteristic feature of these cats was their enlarged upper canines, as seen here in this skull of *Smilodon*.

Christiansen's sample encompassed 24 feline species and 9 fossil sabretooths. He aimed for a comprehensive

analysis of differences in the cranio-mandibular morphology of these specimens by including data on 39 features, so as to offer an evolutionary account that is normalized for body size and set in overall 'morphospace'.

He concludes that the early sabre-toothed cats were not fundamentally distinct from felines, but that different selective pressures came into play as the later (derived) sabretooths evolved. These species set off on an evolutionary trajectory that set a premium on precision killing with large canines. This involved maximizing the extent of their gape, and so a radical reshaping of their skull and mandible. But it came at the expense of a

reduction in bite force.

Such specializations depended on the availability of large prey, and required the predator to be large. By contrast, the superior bite force of the cone-toothed felines allowed a greater variation in their size, as seen in the extant species, and a greater variation in prey size.

These new data, then, add to evidence that — as Christiansen puts it — the derived sabretoothed cats found themselves in an ecological cul-de-sac as ecosystems and climate zones changed.

Tim Lincoln



S. STAMMERS/SPL

the overall *bud6Δ* yeast population is small, the proportion of daughters derived from aged cells is also small. As a result, the average lifespan of the population is increased.

The longer lifespan of *bud6Δ* cells indicates that promoting longevity is not the primary function of the nuclear age barrier. Instead, it may be that, although the barrier ultimately limits population lifespan, it also confers a selective advantage — which is not directly related to ageing — on the daughter cells during the first few cell divisions, when natural selection is strongest. This may occur, for example, by allowing damage-free daughter cells to divide more rapidly or to respond to environmental changes more efficiently. Moreover, it may be that the barrier also functions to modulate asymmetrical inheritance of non-nuclear components. In this regard, it is noteworthy that the same team⁴ has previously described⁹ a similar Bud6- and septin-dependent barrier that limits diffusion of proteins in the endoplasmic reticulum across the bud neck. Whether *bud6Δ* cells are also defective for asymmetrical inheritance of cytoplasmic components remains unknown, and must be explored.

Can Shcheprova and colleagues' observations⁴ inform us about aspects of human ageing — aside from the occasional certainty that our kids are, indeed, making us old before our time? Obvious counterparts of Bud6 and the septin ring do not seem to exist in mammals, and there is no evidence that ERCs contribute to ageing in multicellular organisms. Nonetheless, several genetic factors that modulate replicative ageing in yeast have a similar role

in determining lifespan in the nematode worm *Caenorhabditis elegans*¹⁰. From an evolutionary point of view, yeast and worms are more distantly related than worms and humans, so it is reasonable to speculate that some of these conserved longevity genes also modulate ageing in humans. Moreover, one can imagine that asymmetrical inheritance of damage is likely to occur in certain types of human cell, and that it could be associated with age-related cellular dysfunction and disease. So, as our knowledge of the molecular nature of replicative asymmetry in yeast grows, it is crucial to extend these findings, where possible, to multicellular model systems. Such studies may provide insight into the molecular processes that contribute to ageing and facilitate the development of therapies for various age-related diseases.

Matt Kaeberlein is in the Department of Pathology, University of Washington, Seattle, Washington 98195, USA.
e-mail: kaeber@u.washington.edu

- Steinkraus, K. A., Kaeberlein, M. & Kennedy, B. *Annu. Rev. Cell Dev. Biol.* doi:10.1146/annurev.cellbio.23.090506.123509 (2008).
- Egilmeez, N. K. & Jazwinski, S. M. *J. Bacteriol.* **171**, 37–42 (1989).
- Kennedy, B. K., Austriaco, N. R. Jr & Guarente, L. *J. Cell Biol.* **127**, 1985–1993 (1994).
- Shcheprova, Z., Baldi, S., Buvelot Frei, S., Gonnert, G. & Barral, Y. *Nature* **454**, 728–734 (2008).
- Sinclair, D. A. & Guarente, L. *Cell* **91**, 1033–1042 (1997).
- Aguilaniu, H. *et al. Science* **299**, 1751–1753 (2003).
- Erjavec, N., Larsson, L., Grantham, J. & Nyström, T. *Genes Dev.* **21**, 2410–2421 (2007).
- Lai, C.-Y., Jaruga, E., Borghouts, C. & Jazwinski, S. M. *Genetics* **162**, 73–87 (2002).
- Luedeke, C. *et al. J. Cell Biol.* **169**, 897–908 (2005).
- Smith, E. D. *et al. Genome Res.* **18**, 564–570 (2008).

thought to contribute to ageing in yeast. The best studied of these are extrachromosomal ribosomal-DNA circles (ERCs) — self-replicating plasmids formed by the process of homologous recombination within ribosomal DNA⁵. Consistent with their role as a senescence factor, ERCs are preferentially retained in the mother-cell nucleus and accumulate with age⁵. Mutations that accelerate the rate of ERC formation shorten replicative lifespan, whereas those that reduce ERC accumulation enhance longevity.

Besides ERCs, dysfunctional mitochondria and proteins damaged by reactive oxygen species are actively retained by mother cells and have also been implicated in replicative ageing^{6–8}. Although the importance of such cytoplasmic factors in yeast ageing is less established than that of ERCs, it seems that the accumulation of damage in several mother-cell compartments results in replicative ageing.

Shcheprova *et al.*⁴ advance our understanding of asymmetry between yeast mother and daughter cells, and its importance in ageing, by identifying a diffusion barrier that compartmentalizes the nuclear envelope during budding. This barrier's formation depends on septin proteins, which assemble into a ring around the bud neck and prevent the passage of pre-existing nuclear pores to the daughter cell. Circular DNA molecules (such as ERCs) — and presumably other asymmetrically inherited nuclear factors — are associated with these nuclear-pore complexes, and so are also unable to cross the barrier. This partitioning process requires Bud6, a protein that localizes to the bud neck. Yeast mother cells — and daughter cells — in which the gene encoding this protein has been deleted (*bud6Δ* cells) no longer retain pre-existing nuclear-pore complexes and ERCs, and are consequently long-lived (Fig. 1). This indicates that the presence of the nuclear age barrier limits longevity.

The authors' observations raise some perplexing and intriguing questions. For example, it is not clear why loss of ERCs' asymmetrical distribution increases replicative lifespan. If *bud6Δ* daughter cells inherit ERCs at birth, it seems counterintuitive that they should be long-lived. This apparent paradox can be explained if the ability of an ageing mother cell to pass ERCs to daughter cells more than compensates for the ERCs she herself inherited at birth (Fig. 1). Accordingly, although the daughter cells that arise from young *bud6Δ* mother cells (those that have been through fewer than about eight replicative cycles) inherit some damage at birth, they also go on to continually transfer damage to each of their own daughter cells and, consequently, live longer than do normal cells. Because ERCs and possibly other ageing factors accumulate with age, daughter cells arising from aged *bud6Δ* mother cells (those that have been through more than eight replicative cycles) will inherit a damage burden sufficient to shorten their lifespan. But as the proportion of mother cells of this age or older in

PALAEONTOLOGY

Bite size

Ten thousand or so years ago, the lineage of sabre-toothed cats came to an end. Per Christiansen has revisited the evolutionary history of this distinctive group and that of the other members (felines) of the cat family, with particular reference to changes in the shape of their skull and mandible. He describes his work, which bears on the lively issue of the cause of the sabretooths' demise, in *PLoS ONE* (P. Christiansen *PLoS ONE* **3**, e2807; 2008). It scarcely needs pointing out that the characteristic feature of these cats was their enlarged upper canines, as seen here in this skull of *Smilodon*.

Christiansen's sample encompassed 24 feline species and 9 fossil sabretooths. He aimed for a comprehensive

analysis of differences in the cranio-mandibular morphology of these specimens by including data on 39 features, so as to offer an evolutionary account that is normalized for body size and set in overall 'morphospace'.

He concludes that the early sabre-toothed cats were not fundamentally distinct from felines, but that different selective pressures came into play as the later (derived) sabretooths evolved. These species set off on an evolutionary trajectory that set a premium on precision killing with large canines. This involved maximizing the extent of their gape, and so a radical reshaping of their skull and mandible. But it came at the expense of a

reduction in bite force.

Such specializations depended on the availability of large prey, and required the predator to be large. By contrast, the superior bite force of the cone-toothed felines allowed a greater variation in their size, as seen in the extant species, and a greater variation in prey size.

These new data, then, add to evidence that — as Christiansen puts it — the derived sabretoothed cats found themselves in an ecological cul-de-sac as ecosystems and climate zones changed.

Tim Lincoln



S. STAMMERS/SPL

the overall *bud6Δ* yeast population is small, the proportion of daughters derived from aged cells is also small. As a result, the average lifespan of the population is increased.

The longer lifespan of *bud6Δ* cells indicates that promoting longevity is not the primary function of the nuclear age barrier. Instead, it may be that, although the barrier ultimately limits population lifespan, it also confers a selective advantage — which is not directly related to ageing — on the daughter cells during the first few cell divisions, when natural selection is strongest. This may occur, for example, by allowing damage-free daughter cells to divide more rapidly or to respond to environmental changes more efficiently. Moreover, it may be that the barrier also functions to modulate asymmetrical inheritance of non-nuclear components. In this regard, it is noteworthy that the same team⁴ has previously described⁹ a similar Bud6- and septin-dependent barrier that limits diffusion of proteins in the endoplasmic reticulum across the bud neck. Whether *bud6Δ* cells are also defective for asymmetrical inheritance of cytoplasmic components remains unknown, and must be explored.

Can Shcheprova and colleagues' observations⁴ inform us about aspects of human ageing — aside from the occasional certainty that our kids are, indeed, making us old before our time? Obvious counterparts of Bud6 and the septin ring do not seem to exist in mammals, and there is no evidence that ERCs contribute to ageing in multicellular organisms. Nonetheless, several genetic factors that modulate replicative ageing in yeast have a similar role

in determining lifespan in the nematode worm *Caenorhabditis elegans*¹⁰. From an evolutionary point of view, yeast and worms are more distantly related than worms and humans, so it is reasonable to speculate that some of these conserved longevity genes also modulate ageing in humans. Moreover, one can imagine that asymmetrical inheritance of damage is likely to occur in certain types of human cell, and that it could be associated with age-related cellular dysfunction and disease. So, as our knowledge of the molecular nature of replicative asymmetry in yeast grows, it is crucial to extend these findings, where possible, to multicellular model systems. Such studies may provide insight into the molecular processes that contribute to ageing and facilitate the development of therapies for various age-related diseases.

Matt Kaeberlein is in the Department of Pathology, University of Washington, Seattle, Washington 98195, USA.
e-mail: kaeber@u.washington.edu

- Steinkraus, K. A., Kaeberlein, M. & Kennedy, B. *Annu. Rev. Cell Dev. Biol.* doi:10.1146/annurev.cellbio.23.090506.123509 (2008).
- Egilmeez, N. K. & Jazwinski, S. M. *J. Bacteriol.* **171**, 37–42 (1989).
- Kennedy, B. K., Austriaco, N. R. Jr & Guarente, L. *J. Cell Biol.* **127**, 1985–1993 (1994).
- Shcheprova, Z., Baldi, S., Buvelot Frei, S., Gonnert, G. & Barral, Y. *Nature* **454**, 728–734 (2008).
- Sinclair, D. A. & Guarente, L. *Cell* **91**, 1033–1042 (1997).
- Aguilaniu, H. *et al. Science* **299**, 1751–1753 (2003).
- Erjavec, N., Larsson, L., Grantham, J. & Nyström, T. *Genes Dev.* **21**, 2410–2421 (2007).
- Lai, C.-Y., Jaruga, E., Borghouts, C. & Jazwinski, S. M. *Genetics* **162**, 73–87 (2002).
- Luedeke, C. *et al. J. Cell Biol.* **169**, 897–908 (2005).
- Smith, E. D. *et al. Genome Res.* **18**, 564–570 (2008).

Moving AHEAD with an international human epigenome project

A plan to 'genomicize' epigenomics research and pave the way for breakthroughs in the prevention, diagnosis and treatment of human disease.

The American Association for Cancer Research Human Epigenome Task Force and the European Union, Network of Excellence, Scientific Advisory Board

It is now possible to define whole epigenomes, representing the totality of epigenetic marks in a given cell type. Epigenetic processes are essential for packaging and interpreting the genome, are fundamental to normal development and are increasingly recognized as being involved in human disease. Epigenetic mechanisms include, among other things, histone modification, positioning of histone variants, nucleosome remodelling, DNA methylation, small and non-coding RNAs (Fig. 1). These mechanisms interact with transcription factors and other DNA-binding proteins to regulate gene-expression patterns inherited from cell to cell. The patterns underlie embryonic development, differentiation and cell identity, transitions from a stem cell to a committed cell and responses to environmental signals such as hormones, nutrients, stress and damage.

Although epigenomic changes are heritable in somatic cells, drug treatments could potentially reverse them. This has significant implications for the prevention, diagnosis and treatment of major human diseases and for ageing. Diseases to be targeted could include diabetes, cardiopulmonary diseases, Rett syndrome, other neurological disorders, imprinting disorders, autoimmune diseases and cancer, in which missteps in epigenetic programming have been directly implicated. Indeed, several inhibitors of chromatin-modifying enzymes including histone deacetylase (HDAC) inhibitors and DNA methyltransferase (DNMT) inhibitors have now been approved by the US Food and Drug Administration or are in clinical trials with good prognosis for tumour regression. Epigenetic therapy is now a reality, but to maximize the potential of such therapeutic approaches, it is crucial that there be a more comprehensive characterization of the epigenetic changes that occur during normal development, adult cell renewal and disease, and of the relationships between genetic and epigenetic variation and their impact on health.

The time is right for a major effort to decode

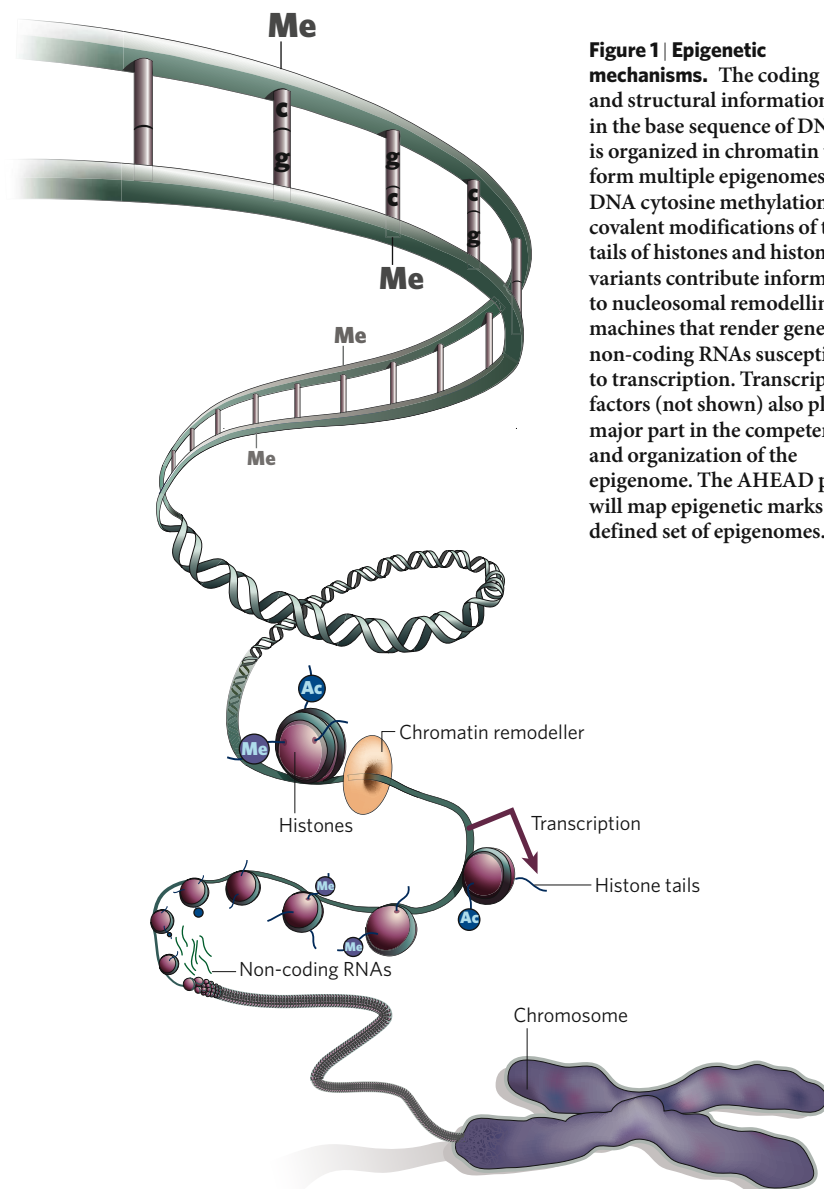


Figure 1 | Epigenetic mechanisms. The coding and structural information in the base sequence of DNA is organized in chromatin to form multiple epigenomes. DNA cytosine methylation and covalent modifications of the tails of histones and histone variants contribute information to nucleosomal remodelling machines that render genes and non-coding RNAs susceptible to transcription. Transcription factors (not shown) also play a major part in the competence and organization of the epigenome. The AHEAD project will map epigenetic marks in a defined set of epigenomes.

the human epigenome, and we urge the community to join in a coordinated effort in support of the Alliance for the Human Epigenome and Disease (AHEAD) to help solve the problems of cancer and other intractable diseases. Just as the Human Genome Project provided a reference 'normal' sequence for studying human disease, the goal of the

AHEAD project is to provide high-resolution reference epigenome maps. These maps would be of great use in basic and applied research, would have an immediate impact on understanding many diseases, and would lead to the discovery of new means to control the diseases. Although the project should have a human focus, it will be essential to use model

Box 1 | A European success story

European researchers in the late 1990s were quick to take up the fast-emerging field of epigenetics, building on early discoveries in X-chromosome inactivation, Polycomb regulation, dosage compensation, imprinting, heterochromatin, DNA methylation, nucleosome modification and RNA interference (RNAi). The Epigenome Network of Excellence (NoE) was created by the European Commission in 2004 with the long-term aim of tackling major epigenetic questions in the post-genomic era. Under the direction of Thomas Jenuwein of the Research Institute of Molecular Pathology in Vienna, assisted by Phil Avner of the Pasteur Institute in Paris and Geneviève Almouzni of the Curie Institute in Paris, the intention was to create a durable structure for epigenetics research in Europe, promoting collaborations, sharing technologies and resources, and training and fostering the development of leaders in the field.

Now, after four of the five years planned, the NoE includes an original core of 25 members plus 35 associate members and 23 Newly Established Teams (NETs): a total of 83 laboratories from 12 European countries,

as well as 350 laboratories worldwide that participate in ongoing activities through the NoE website. Expansion of the NoE resulted in the creation of Associate Members and, most importantly, the NETs. Through the NET programme, talented young European investigators in their first four years of independent research are selected to participate and provided with moderate financial support. The NET programme has been a resounding success and has given a powerful impetus to a cohort of young research talent.

The NoE programme of activities has promoted a common identity and generated numerous collaborations through training and mobility programmes, workshops, meetings and dissemination of technologies. In 2007 alone, the NoE collectively published some 300 research articles.

A major effort has also been made to reach out to the broader community with a scientific website (www.epigenome-noe.net), and to the general public, through conferences, newspaper articles, and a public website (www.epigenome.eu), where news, information, explanations of the concepts, significance and applications of epigenetic research



in lay terms are provided in seven languages. Several NoE members are involved in biotechnology companies and technology transfer has been the subject of NoE workshops.

These achievements have resulted from a conjunction of circumstance, hard work and the talent assembled by the NoE. Europe is in many respects pre-eminent in the field of epigenetics and the NoE could serve as a model for similar programmes worldwide. But, as its five-year mandate approaches its end, the future of the NoE remains uncertain. The organization, coordination and

momentum achieved should be nurtured, supported and exploited to broaden connections with biotechnology, agriculture and particularly with the biochemical and pharmaceutical industries that are such key components of the European technological landscape. With the worldwide interest in epigenetics, this seems an inopportune moment for the European Commission to reduce support for this research and risk losing momentum and brainpower. On the contrary, the opportunity should be seized to forge close links with international efforts such as AHEAD.

organisms to obtain mechanistic insights as to the functionality of epigenomic parameters or 'codes'. Studying, genome-wide, the increasing number of interacting epigenetic mechanisms in a spectrum of cell types will involve data generation on a massive scale. An international project would provide the bioinformatics infrastructure needed to ensure that epigenomics research can be integrated with genomic data, and be utilized efficiently to advance the knowledge of human health and disease. Here, we discuss the benefits of the AHEAD framework to coordinate and plan an international Human Epigenome Project.

Early steps

Individual investigators have been studying epigenetics for several decades; however, concerted efforts to organize the epigenomics research community are quite recent, and none has sought to engage the community on a broad-based, international level.

Europe has a strong tradition for epigenetics research that has been recognized by European Union funding programmes and by individual national initiatives. More than €50 million (US\$79 million) has been allocated to networks and consortia that focus on central epigenetic questions such as DNA methylation (HEP, Human Epigenome

Project), chromatin profiling (HEROIC, High-Throughput Epigenetic Regulatory Organization In Chromatin), and treatment of neoplastic disease (EPITRON, Epigenetic Treatment Of Neoplastic Disease). A special function is provided by the Epigenome Network of Excellence (NoE), created by the European Commission in 2004 (see Box 1).

In the past few years, there have been several efforts to organize the epigenetics research community in the United States and develop the support and structure for a human epigenome project. A 2004 National Cancer Institute (NCI)-sponsored Epigenetic Mechanisms in Cancer Think Tank concluded that the development of a US human epigenome project of analogous scope to the Human Genome Project should be of the highest priority (<http://www.cancer.gov/think-tanks-cancer-biology/page7>). This was followed by an NCI workshop in 2005 attended by programme staff from many National Institutes of Health (NIH) institutes. Key recommendations were: 1) comprehensive analysis of reference epigenomes, focused on stem cells and key differentiated lineages that can be modelled experimentally; 2) development of a standardized set of reagents, including monoclonal antibodies and common tissue substrates; and 3) development of

a computational infrastructure for data sharing and display that puts epigenetic data in the hands of non-epigenetic investigators¹.

The American Association for Cancer Research (AACR) organized a Human Epigenome Workshop in June 2005 and also a follow-up workshop in July 2006 that focused on planning an international project to map a defined number of human epigenomes (<http://www.aacr.org/page9673.aspx>)². The consensus of workshop participants was that there are compelling reasons from both scientific and public-health perspectives to initiate a human epigenome project that could take full advantage of advances in several existing US and European initiatives.

On the heels of these workshops, the AACR Human Epigenome Task Force, a cross-disciplinary group of international investigators, was formed to design a strategy and develop a timetable for the implementation of an international Human Epigenome Project. As described below, the task force recommends the formation of AHEAD to coordinate a transdisciplinary, international project to map a defined subset of robust epigenetic markers in a limited number of human tissues at different stages of development and to develop a bioinformatics infrastructure to support the collection of epigenomic data. The efforts

Table 1 | Potential reference cell systems for epigenomic analysis

Organ system	Tissue progenitor	Differentiated cell	Altered disease state
Embryo	Embryonic stem cells		
Haematopoietic	CD34+	Polymorphonuclear leukocytes, lymphoid, etc.	Acute myelogenous leukemia, Acute lymphoblastic leukemia, Chronic myelogenous leukemia, Myelodysplastic syndrome
Breast	CD24+and/or CD44+	Myoepithelial, luminal, fibroblasts	Breast cancer
Liver	Oval cell	Hepatocytes, stromal cells	Cirrhosis, viral disease, Hepatocellular carcinoma

of this group will surely be bolstered by the recent and exciting decision of the NIH to fund epigenomics research as an interagency Road-map initiative.

Asia, too, has been active in fostering epigenomics research, with a major emphasis placed on disease epigenomes, especially those in liver and gastric cancers. An international meeting, Genome-wide Epigenetics 2005, was held in Tokyo³. Scientists from Yonsei University (South Korea), the National Cancer Center (Japan), the Shanghai Cancer Institute (China) and the Genome Institute (Singapore) also organized meetings to facilitate the exchange of information in epigenomics and held their first meeting in Seoul in 2006, followed by another conference in Osaka in 2007. In December 2006, a Japanese Society for Epigenetics was formed. Clearly, Asia is now poised to contribute strongly to global epigenomics research.

National support for the Human Epigenome Project is also mounting in Australia with the formation of the Australian Alliance for Epigenetics to begin in late 2008. Australian meetings devoted to epigenetics were initiated in 1996 in Heron Island with Susan Clark of the Garvan Institute of Medical Research, Sydney, holding the first workshop on bisulphite sequencing and this has been followed by biannual meetings, the most recent being held in Perth in November 2007 to showcase Australia's strengths in the global epigenetic research arena.

The scope of AHEAD

AHEAD would aim to provide reference epigenomes for key cellular states, such

as a stem-cell phenotype, proliferation, differentiation, senescence and stress, using common core specimens from different cell types, and to develop a bioinformatics infrastructure to support the collection and integration of epigenomic data. The precise objectives of such a project still need to be delineated. However, we would almost certainly expect AHEAD to 1) provide complete epigenome maps at very high resolution for important histone modifications across diverse cellular states in both human and mouse; 2) to complete and catalogue epigenome maps of model yeasts (*Saccharomyces cerevisiae* and *S. pombe*), plants (*Arabidopsis* and rice), and animals (*Caenorhabditis elegans* and *Drosophila*); 3) to deliver a high resolution DNA methylation map of the entire human genome in defined cell types and a landmark map for transcription start sites of all protein coding genes and a representative number of other features throughout the genome; 4) define non-coding and small RNAs; and 5) to establish a bioinformatics platform including a relational database, website and suite of analytical tools to organize, integrate and display whole epigenomic data on model organisms and humans. AHEAD will develop standardized procedures that assure data quality in the choice of reagents (antibodies), the experimental procedures (a minimal set of biological replicates) and data analysis (appropriate statistical procedures).

AHEAD would differ from and complement other ongoing projects such as ENCODE (ENCyclopedia Of DNA Elements). The ENCODE project is focused on defining the functional sequences in the genome, whereas AHEAD would define the patterns of epigenetic regulation occurring at those sequences in different cell states. The potential synergies between these two projects in terms of development of innovative molecular and computational techniques would enhance each of them. It would be valuable to coordinate such individual projects in the interests of efficiency and potential scientific insights. Some of these objectives are exemplified by the current efforts on model organisms carried out by the modENCODE (Model Organism ENCyclopedia Of DNA Elements) programme.

Reference epigenomes

In a multicellular organism, there are many potential epigenomes that define each cell type

and reflect the current and past environments of individual cells. Epigenetic programming may even precede lineage choice and amplify signals from the environment. A major goal of AHEAD would be to map the epigenome in normal tissues and differentiating cells so that they could be compared with disease states, including but not restricted to cancer, that arise in these tissues. Therefore, the selection of which cell types to use and which epigenetic modifications to examine deserves careful consideration. Many environmental factors, such as ageing, hormonal milieu, diet and infection must be taken into account. Emphasis should be placed on comparing embryonic stem cells, adult stem and precursor cells, and related differentiated cells of primary human tissues, with the diseased state.

The reference systems selected should meet several criteria: 1) cells should be easy to sample in a reproducible fashion (renewable if possible); 2) cell numbers should be sufficient for analyses of DNA methylation and chromatin modifications; 3) cell progenitors should be identified that can be suitably manipulated and harvested in a pure state; 4) where possible, cells should be amenable to tissue reconstruction and three-dimensional model systems; and, of course, 5) systems must provide insight into key differentiation and related disease states. Some examples of suitable systems are given in Table 1.

Specific modifications to histones often correlate with gene activation or repression — for example lysine acetylation and trimethylation of lysine 4 of histone H3 (H3K4me3) are permissive for gene activation whereas H3K9me2 and H3K27me3 correlate with transcriptional silencing. Often, activating and repressive histone marks co-exist at gene start sites, reflecting perhaps epigenetic heterogeneity among otherwise similar cells, and it is the balance between these marks that determines gene expression states^{4–6}. Because there are so many chromatin modifications, it may prove challenging to examine all of them in every cell type. We recommend that the initial set of reference epigenomes focus on several of the better understood covalent markers, such as those listed in Table 2, that could be mapped in a 'first pass' project defining reference epigenomes.

Several high-profile studies, including the human ENCODE maps⁷ and genome-wide

Box 2 | Heritability

Epigenetics is often defined as somatically heritable changes in gene expression that do not involve changes in base sequence. DNA methylation patterns are known to be heritable through S phase of the cell cycle, however it has been more difficult to demonstrate mechanisms for the heritability of other chromatin modifications that will be mapped. Also, the abilities of transcription factors to reprogram the epigenome in normal and experimental situations need to be considered. The AHEAD project will help illuminate processes of heritability and reprogramming by providing genome-wide reference maps.

Table 2 | Recommended histone markers with known function to define reference epigenomes

Histone modification	Function
H3K9ac, H4K16ac and H3K4me	Transcriptional 'ON' states
H3K9me and H3K27me	Transcriptional 'OFF' states
H3K36me	Elongation; primary transcripts
H3K9me and HP1	Chromosome organization
H4K20me and H2B514ph	Stress/damage response

profiling for histone modifications in mouse and human cells^{8,9}, have been published recently. Although they present valuable first-generation epigenetic maps, these efforts largely focused on transcriptional regulation and functional sequences. We still need to know more about epigenetic control and plasticity in intergenic regions, mechanisms of heritability (Box 2), the role of repetitive elements, non-coding and small RNAs, and finally how epigenetic marking contributes to chromosome segregation, stress response and transducing signals from the environment.

Advances in technology

Our understanding of the epigenome has been transformed in recent years by a succession of technological innovations. Approaches involving microarrays and, most recently, ultra-high-throughput sequencing technology have been applied to map chromatin modifications, cytosine methylation and non-coding RNAs across chromosomes and even entire genomes.

Genome-scale studies of histone modifications and other aspects of chromatin structure have typically relied on an immunological procedure, chromatin immunoprecipitation, in which specific antibodies are used to enrich chromatin. For example, an antibody against histone H3 acetylated at lysine 9 can be used to isolate genomic regions associated with this activating modification. Isolated DNA is then interrogated by microarray hybridization or by deep sequencing. Microarrays are a well-established readout, particularly suited to studies that require high coverage of a small subset of a given genome (for example, annotated gene promoters). Resolution on the order of the nucleosome (a few hundred bases) can be achieved with sufficiently dense tiling arrays. Recent studies — several dozen whole mammalian genome data sets have already been reported^{8,9} — have leveraged emerging ultra-high-throughput sequencing technology to ‘deep-sequence’ chromatin-immunoprecipitated DNA to generate genome-wide chromatin state maps. This sequencing approach offers potential advantages over arrays in precision, throughput and genome coverage. Notably, sequencing requires orders-of-magnitude less DNA than microarrays, and this enhanced sensitivity should enable studies of primary tissues, disease samples, and other limited cell populations of high biological importance.

The quality of the data derived from studies relying on chromatin immunoprecipitation depends mainly on the quality of the antibodies used, and there is an essential need for high-titre polyclonal antibodies with high specificity. Standardization of monoclonal antibodies has proven difficult because of low avidity and only limited epitope recognition. Reference maps defined by AHEAD, as well as those already developed from studies in human cells and model organisms, would provide a platform to assess the quality and

robustness of modification-specific antibodies.

Rapid progress has also been made towards characterizing the global distribution of cytosine methylation. Genome-scale assays for detecting this epigenetic modification fall into two general categories. The gold-standard technique for reading out the methylation state of individual cytosines is bisulphite sequencing developed by Clark and Marianne Frommer at the University of Sydney¹⁰, in which unmethylated cytosines are converted to uracils and read as thymine, while methylated ones are protected from conversion. Although this method yields precise nucleotide-resolution data, bisulphite sequencing is challenging to scale — regions of interest must be sequenced several times in order for methylation patterns in a given cell type to be appreciated. A recent study showed the potential of combining bisulphite methodology with ultra-high-throughput sequencing by deciphering the entire *Arabidopsis* ‘DNA methylome’ at nucleotide resolution¹¹. Still, bisulphite studies of the much larger human methylome have thus far been limited to relatively small subsets of genome¹². For more comprehensive coverage, many investigators have turned to alternative approaches that involve the isolation of methylated (or unmethylated) fractions of genome by methylation-sensitive restriction or immunoprecipitation with a methylcytosine-specific antibody. The isolated fractions are then interrogated on microarrays. Several successes have been described recently using this genome fractionation/microarray approach, including a profile of promoter methylation in human fibroblasts¹³ and drafts of the complete *Arabidopsis* methylome^{14–16}.

Model organisms

Model organisms have led the way in understanding epigenetic mechanisms of gene regulation. Position effect variegation, imprinting, transposon silencing, DNA methylation and histone modification were all discovered in model organisms and later found in humans. RNA interference and its role in epigenetics were also first described in model organisms (plants and worms).

The development of infrastructure to study model-organism epigenomes is under way and is having a major impact on human epigenome research, with hundreds of whole-genome or whole-chromosome profiles of histone modifications and DNA methylation already published in yeast, *Arabidopsis*, *Drosophila* and the mouse. Importantly, mutants in epigenetic modification, gene silencing, development and metabolism, as well as disease models, can be readily obtained and manipulated, and will allow the epigenetic pathways that are responsible for controlling genomic output to be determined first in model organisms. Other tools such as lineage-specific green fluorescent protein reporters and tagged histones have been engineered in mouse and *Arabidopsis*,

allowing cell-type specific epigenomic profiles to be determined through cell sorting or immunoprecipitation. It is anticipated that new technologies such as massively parallel sequencing will make the generation of these profiles so straightforward that the limitations will be in the data analysis and the ability to take advantage of such genetic and cell biological tools. Reference epigenomes for several model organisms including *S. cerevisiae*, *Arabidopsis*, *Drosophila*, *C. elegans*, mouse and others will form an integral part of AHEAD. Clearly, AHEAD would enable coordination and expansion of these projects and rapid translation into human studies.

Computational challenges

As a comprehensive human epigenome project will require a strong bioinformatics platform, an early priority of AHEAD would be to establish a central relational database and a web interface to present data to the scientific community. This website would include analytical and statistical tools that could be used dynamically to process data for visualization. It would have tremendous utility for investigators, allowing immediate access to detailed maps for a locus of interest in the same way that the Human Genome Project provides sequence data. By initially focusing on establishing data interoperability, it is hoped that the database, or later versions of it, would allow for the simultaneous display of integrated epigenomic parameters on the entire human genome. In addition, not yet addressed is how the complex system of epigenomic regulation should be treated as a whole: the data could be a powerful resource for the emerging field of systems biology, allowing insights into the stability of the epigenomic networks and how they can be perturbed in disease states.

The development of these resources would be coordinated with parallel efforts already under way as part of the caBIG (the NIH Cancer Biomedical Informatics Grid) and Cancer Genome Atlas initiatives, among others. These resources could be distributed through the MGED (Microarray Gene Expression Data Society) and the BioConductor (open-source software) organizations. In addition, one could build on the sophisticated approaches of existing genome browsers used for data access and visualization, such as those at University of California, Santa Cruz, and ENSEMBL, to develop innovative means of data representation and web-based data analysis that would be of immense value to and foster collaborations among the scientific community.

Summary and perspectives

Dramatic changes in technologies have made AHEAD eminently feasible. Successful coordination of this multifaceted project through an AHEAD International Steering Committee would be essential if rapid progress is to be made, and it will require leadership from many

stakeholders at all levels. The major challenges of AHEAD include the initial selection of the most important epigenomic markers, the identification of appropriate, pure cell populations, and the handling and presentation of the data. None of these challenges is insurmountable, and there are enormous potential scientific and public health benefits.

AHEAD would effectively integrate epigenetics research that is currently being conducted on a piecemeal basis around the world and would pave the way for breakthroughs in understanding normal and pathological processes. This research would offer significant scientific opportunities to maximize translational research to prevent and cure human diseases. The selection of epigenetics as a new NIH Roadmap initiative in the United States and the European Union's Network of Excellence

are important in helping establish the infrastructure needed to advance the field. We must now move AHEAD in earnest to achieve the goals of an international Human Epigenome Project. ■

1. Feinberg, A. P., Jones, P. A. & Ault, G. National Cancer Institute (NCI), Division of Cancer Biology (DCB) Workshop on Defining the Epigenome: Addressing the Value and Scope of a Human Epigenome Project. <http://dcb.nci.nih.gov/Workshoprpt.cfm>.
2. Jones, P. A. & Martienssen, R. *Cancer Res.* **65**, 11241–11246 (2005).
3. Lieb, J. D. *et al. Cytogenet. Genome Res.* **114**, 1–15 (2006).
4. Azuara, V. *et al. Nature Cell Biol.* **8**, 532–538 (2006).
5. Bernstein, B. E. *et al. Cell* **125**, 315–326 (2006).
6. Guenther, M. G., Levine, S. S., Boyer, L. A., Jaenisch, R. & Young, R. A. *Cell* **130**, 77–88 (2007).
7. Birney, E. *et al. Nature* **447**, 799–816 (2007).
8. Barski, A. *et al. Cell* **129**, 823–837 (2007).
9. Mikkelsen, T. S. *et al. Nature* **448**, 553–560 (2007).
10. Clark, S. J., Harrison, J., Paul, C. L. & Frommer, M. F. *Nucleic Acids Res.* **22**, 2990–2997 (1994).

11. Cokus, S. J. *et al. Nature* **452**, 215–219 (2008).
12. Eckhardt, F. *et al. Nature Genet.* **38**, 1378–1385 (2006).
13. Weber, M. *et al. Nature Genet.* **39**, 457–466 (2007).
14. Zhang, X. *et al. Cell* **126**, 1189–1201 (2006).
15. Zilberman, D., Gehring, M., Tran, R. K., Ballinger, T. & Henikoff, S. *Nature Genet.* **39**, 61–69 (2007).
16. Vaughn, M. W. *et al. PLoS Biol.* **5**, e174 (2007).

Acknowledgements We gratefully acknowledge Kimberly Sabelko from the AACR for her assistance in editing this manuscript. We also thank the AACR and Margaret Foti for supporting and funding the establishment of the Human Epigenome Task Force.

Author contributions P.A.J., S.B.B., B.E.B., A.P.F., J.M.G., T.J., R.M., T.U., and V.P., C.D.A., S.C.E., J.R. and C.W. all contributed to this manuscript.

Author Information Correspondence and requests for materials should be addressed to P.A.J. (e-mail: jones_p@ccnt.hsc.usc.edu).

See Tech Feature, page 795.

The American Association for Cancer Research Human Epigenome Task Force

Peter A. Jones¹, Trevor K. Archer², Stephen B. Baylin³, Stephan Beck⁴, Shelley Berger⁵, Bradley E. Bernstein⁶, John D. Carpten⁷, Susan J. Clark⁸, Joseph F. Costello⁹, Rebecca W. Doerge¹⁰, Manel Esteller¹¹, Andrew P. Feinberg¹², Thomas R. Gingeras¹³, John M. Greally¹⁴, Steven Henikoff¹⁵, James G. Herman³, Laurie Jackson-Grusby¹⁶, Thomas Jenuwein¹⁷, Randy L. Jirtle¹⁸, Young-Joon Kim¹⁹, Peter W. Laird¹, Bing Lim²⁰, Robert Martienssen²¹, Kornelia Polyak²², Henk Stunnenberg²³, Thea Dorothy Tlsty²⁴, Benjamin Tycko²⁵, Toshikazu Ushijima²⁶ and Jingde Zhu²⁷

The European Union, Network of Excellence, Scientific Advisory Board

Vincenzo Pirrotta²⁸, C. David Allis²⁹, Sarah C. Elgin³⁰, Peter A. Jones¹, Robert Martienssen²¹, Jasper Rine³¹ and Carl Wu³²

¹USC/Norris Comprehensive Cancer Center, Los Angeles, California 90089, USA. ²NIH-NIEHS, Research Triangle Park, North Carolina 27709, USA. ³Johns Hopkins University, Baltimore, Maryland 21231, USA. ⁴Wellcome Trust Sanger Institute, Hinxton, Cambridge CB10 1SA, UK. ⁵The Wistar Institute, Philadelphia, Pennsylvania 19104, USA. ⁶Massachusetts General Hospital, Charlestown, Massachusetts 02129, USA. ⁷Gen, Phoenix, Arizona 85004, USA. ⁸Garvan Institute of Medical Research, Sydney, NSW 2010, Australia. ⁹Department of Neurological Surgery, University of California, San Francisco, San Francisco, California 94143, USA. ¹⁰Purdue University, West Lafayette, Indiana 47907, USA. ¹¹Spanish National Cancer Center, Madrid, Spain. ¹²Departments of Medicine, Oncology and Molecular Biology & Genetics, Johns Hopkins University School of Medicine, Baltimore, Maryland 21205, USA. ¹³Affymetrix, Santa Clara, California, 95051, USA. ¹⁴Albert Einstein College of Medicine, Bronx, New York 10461, USA. ¹⁵Fred Hutchinson Cancer Research Center, Seattle, Washington 98104, USA. ¹⁶Harvard Medical School, Boston, Massachusetts 02115, USA. ¹⁷Research Institute of Molecular Pathology, Vienna 1030, Austria. ¹⁸Duke University Medical Center, Durham, North Carolina 27710, USA. ¹⁹Yonsei University, Seoul 120-719, Korea. ²⁰Genome Institute of Singapore, Singapore 138672. ²¹Cold Spring Harbor Laboratory, Cold Spring Harbor, New York 11724, USA. ²²Dana-Farber Cancer Institute, Boston, Massachusetts 02115, USA. ²³Radboud University, Nijmegen 6525, the Netherlands. ²⁴Department of Pathology, University of California, San Francisco, San Francisco, California, 94143, USA. ²⁵Columbia University, New York, New York 10032, USA. ²⁶National Cancer Center Research Institute, Tokyo 104-0045, Japan. ²⁷Shanghai Cancer Institute, Shanghai 200032, China. ²⁸Rutgers University, Piscataway, New Jersey 08854, USA. ²⁹Rockefeller University, New York, New York 10021, USA. ³⁰Washington University, Saint Louis, Missouri 63130, USA. ³¹University of California Berkeley, Berkeley, California 94720-3202, USA. ³²National Cancer Institute, Bethesda, Maryland 20892, USA.

Secreted transcription factor controls *Mycobacterium tuberculosis* virulence

Sridharan Raghavan¹, Paolo Manzanillo^{1*}, Kaman Chan^{1*}, Cole Dovey¹ & Jeffery S. Cox¹

Bacterial pathogens trigger specialized virulence factor secretion systems on encountering host cells. The ESX-1 protein secretion system of *Mycobacterium tuberculosis*—the causative agent of the human disease tuberculosis—delivers bacterial proteins into host cells during infection and is critical for virulence, but how it is regulated is unknown. Here we show that EspR (also known as Rv3849) is a key regulator of ESX-1 that is required for secretion and virulence in mice. EspR activates transcription of an operon that includes three ESX-1 components, *Rv3616c–Rv3614c*, whose expression in turn promotes secretion of ESX-1 substrates. EspR directly binds to and activates the *Rv3616c–Rv3614c* promoter and, unexpectedly, is itself secreted from the bacterial cell by the ESX-1 system that it regulates. Efflux of the DNA-binding regulator results in reduced *Rv3616c–Rv3614c* transcription, and thus reduced ESX-1 secretion. Our results reveal a direct negative feedback loop that regulates the activity of a secretion system essential for virulence. As the virulence factors secreted by the ESX-1 system are highly antigenic, fine control of secretion may be critical to successful infection.

Pathogenic microbes sense their environment and change their physiology to interact with the host. *Mycobacterium tuberculosis*, a pathogen of global importance, utilizes the ESX-1 protein secretion system to export virulence factors that disarm host macrophages^{1–3}. ESX-1, also termed type VII secretion, is critical for virulence, and transports proteins from inside the bacterium across the cell envelope^{1,2,4,5}. The entire ESX-1 system has been implicated in innate immune modulation, especially early after infection in macrophages^{1,6–8}. Consistent with this view, ESX-1 mutants are specifically defective in the early stages of growth in mice^{1,2,5}. However, the role of the major secreted substrates, ESAT-6 and CFP-10, in virulence is a matter of current debate^{2,3,9,10}. Furthermore, these proteins are also potent antigens that elicit protection against tuberculosis in animal models, and are important components of vaccines currently in clinical trials^{11–15}.

The core machinery of the ESX-1 pathway, as well as ESAT-6 and CFP-10, are encoded at the genomic locus known as region of difference 1 (RD1), which is absent in the BCG vaccine strain¹⁶. Many other proteins encoded at the RD1 locus are also required for ESX-1 function, suggesting that this secretion apparatus is complex^{17–19}.

In addition to RD1-encoded genes, a separate locus required for ESX-1 system function has also been identified^{20,21}. This locus contains the genes *Rv3616c–Rv3614c* that are required for ESX-1 system activity; the *Rv3616c* protein, termed EspA, is a secreted substrate of the pathway^{20,21}. The discovery of this additional substrate also revealed the surprising phenomenon of mutually dependent export, in which the secretion of each substrate relies upon the secretion of the other substrates²⁰. Although the mechanism underlying this co-dependence is not understood, it raises the possibility that substrates exert control over the secretion activity of the entire ESX-1 pathway. The phenomenon of mutually dependent secretion has complicated the genetic dissection of effector functions of single ESX-1 substrates, and the specific functions of the known ESX-1 substrates remain unclear.

In this work, we have identified a novel secreted substrate of the ESX-1 pathway, EspR. Surprisingly, this protein is a DNA-binding

transcription factor that regulates *M. tuberculosis* genes that are required for ESX-1. The finding of a secreted DNA-binding protein reveals a novel homeostatic feedback mechanism by which the activity of a secretion system is monitored and regulated through coupling between secretion and transcription.

EspR (Rv3849) is secreted by ESX-1

In a genetic screen, we identified an *M. tuberculosis* transposon mutant that induced high levels of IL-12 from macrophages, a common phenotype of ESX-1 secretion mutants¹, and determined that the transposable element inserted 13 nucleotides upstream of the gene *Rv3849*. Immunoblotting proteins from bacterial pellets and culture supernatants demonstrated that Rv3849 was secreted from wild-type cells, and although protein levels were severely reduced in the *Rv3849*[−] mutant, a low level of Rv3849 was still present in cell lysates (Fig. 1a, lanes 1–4). Notably, ESAT-6 secretion was severely diminished in the *Rv3849*[−] mutant, with the block nearly equivalent to that of a mutant lacking the putative pore protein Rv3877 (Fig. 1a, lanes 1–4, 7 and 8). This defect was corrected upon integration of a single copy of the *Rv3849* gene into mutant bacteria (Fig. 1a, lanes 5 and 6). Importantly, Rv3849 secretion is blocked in bacteria lacking either Rv3877 (lanes 7 and 8) or ESAT-6 (encoded by *esxA*, lanes 9 and 10). Like other ESX-1 secretion mutants¹, *Rv3849*[−] mutant cells exhibited a characteristic growth defect during infection of mice (Fig. 1b). Mutant bacteria grew poorly in the lungs during the first five days post-infection but this was followed by a period in which the viable mutants recovered normal growth kinetics. Although the mutant cells established a stable infection by three weeks, the bacterial burden was approximately an order of magnitude less than the wild type (Fig. 1b). Therefore, Rv3849 is a new ESX-1 substrate that, like ESAT-6, CFP-10 and EspA, is required for the function of the entire ESX-1 system²⁰.

EspR is required for ESX-1 gene expression

Although standard homology and motif searches failed to identify any putative functional domains in Rv3849, a secondary structure-based

¹Department of Microbiology and Immunology, Program in Microbial Pathogenesis and Host Defense, University of California, San Francisco, 600 16th Street, Campus Box 2200, San Francisco, California 94143-2200, USA.

*These authors contributed equally to this work.

homology search²² identified a significant match to the *Bacillus subtilis* transcription factor SinR, a helix–turn–helix (HTH) DNA-binding protein²³. Modelling the three-dimensional structure of Rv3849 using the crystal structure of SinR as a template revealed considerable similarity between the two proteins, especially at the amino-terminal HTH domain (Fig. 1c). Thus, although Rv3849 is an ESX-1 secreted protein, it bears striking similarities to a DNA-binding protein.

To determine if Rv3849 is a transcriptional regulator, we used *M. tuberculosis* microarrays to compare the global transcriptional profile between wild-type and *Rv3849*[−] bacteria grown in culture. Of the 4,505 genes represented on the array, expression of only five operons, representing 14 genes, was altered in *Rv3849*[−] mutant cells (Fig. 1d, Supplementary Table 1). Three operons were activated in the *Rv3849*[−] mutant, including a putative operon containing the genes *MT2035*, *Rv1982c* and *nrdF1* (Fig. 1d). Notably, the *Rv3616c*–*Rv3612c* operon was one of two operons that were significantly downregulated in the *Rv3849*[−] mutant (Fig. 1d). Quantitative RT–PCR (qPCR) to measure *nrdF1* and *Rv3615c* messenger RNA levels provided additional support for the microarray data (Supplementary Fig. 1). Because our *Rv3849* mutant contains a transposon insertion in the promoter but retains an intact open reading frame, we made a deletion mutant in which the entire *Rv3849* gene was replaced by allelic exchange. Importantly, the secretion and transcriptional phenotypes of the transposon mutant are indistinguishable from those of the deletion mutant (Supplementary Fig. 2). Because the genes *Rv3616c*, *Rv3615c* and *Rv3614c* are required for ESX-1, this result provided a direct link between Rv3849 and the secretion system. Overall, depletion of Rv3849 led to both positive and negative effects on the expression of a small number of *M. tuberculosis* genes, including one locus critical for ESX-1 function. Given the regulatory function of Rv3849 in ESX-1 secretion, we renamed this protein EspR (ESX-1 secreted protein regulator).

As ESX-1 functions early during infection^{1,8}, we reasoned that EspR may activate ESX-1 secretion upon phagocytosis by macrophages. To assay EspR activity, we measured transcription of the *Rv3616c*–*Rv3612c* operon during infection by qPCR. *Rv3616c*–*Rv3612c* expression was significantly induced by two hours post-infection in wild-type bacteria (Fig. 1e), consistent with the recent finding that this operon is among a set of genes activated upon phagosomal acidification²⁴. In *espR*[−] bacteria, however, *Rv3616c*–*Rv3612c* mRNA levels were near background two hours after infection, and although some induction occurred by four hours, the levels were much reduced compared to wild-type *M. tuberculosis* (Fig. 1e). These data suggest that EspR is induced upon phagocytosis, activating expression of downstream ESX-1 components.

EspR binds DNA and activates transcription

To determine if EspR directly regulates *Rv3616c*–*Rv3612c*, we measured the activity of an *Rv3616c* promoter–*lacZ* transcriptional reporter in *Mycobacterium smegmatis*, a non-pathogenic relative of *M. tuberculosis* that lacks the *Rv3616c*–*Rv3612c* operon but encodes a homologue of EspR. The 680-base-pair (680-bp) promoter induced slight *lacZ* expression compared to the promoter-less construct, but was strongly activated by expression of *M. tuberculosis* EspR (Fig. 2a). Deletions from the 5′ end of the promoter indicated that sequences between 427 and 520 bp upstream of *Rv3616c* were required for EspR activity (Fig. 2a). In electrophoretic mobility shift assays (EMSA), purified EspR, expressed as an N-terminal fusion with maltose binding protein (MBP–EspR), bound directly to probes that included at least 520 bp of the *Rv3616c* promoter (Fig. 2b). The 427-bp probe was weakly bound by the fusion protein (Fig. 2b), and no binding was detectable with a 320-bp probe (data not shown). Thus, EspR binds directly to the *Rv3616c* promoter and activates transcription.

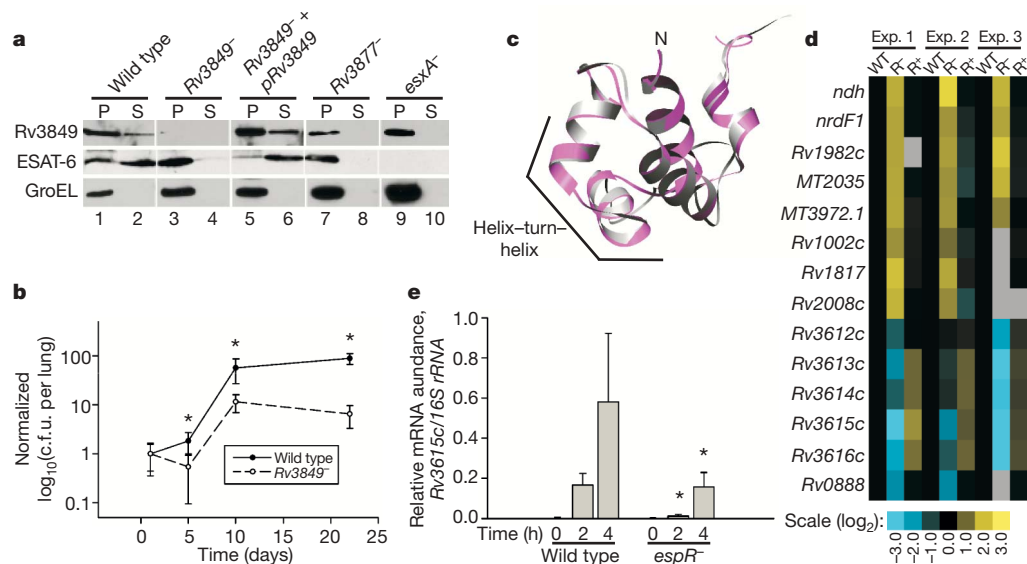


Figure 1 | Rv3849 is a secreted regulator of the ESX-1 system. **a**, Cell pellets (P) and culture supernatants (S) from wild-type, *Rv3849*[−], *Rv3849*[−] complemented, *Rv3877*[−], and *esxA*[−] bacteria were probed for Rv3849, ESAT-6 and GroEL by western blot. GroEL, an intracellular protein used as a lysis control, was detected in cell pellets of all strains and was absent from supernatants. **b**, C57BL/6 mice were infected with 1×10^6 colony-forming units (c.f.u.) of wild-type and *Rv3849*[−] mutant *M. tuberculosis* through the intravenous route. Bacterial growth was monitored by counting c.f.u. from lung homogenates, and values at each time point were normalized to the initial dose at 24 h post-infection. For each bacterial strain, mean ± s.d. is shown from five mice at each time point. **P* < 0.05 between *Rv3849*[−] and wild-type. **c**, The crystal structure of SinR (violet) was used as a template to generate a structural model of Rv3849 (silver) using the program Modeller. The overlay of the N-terminal HTH domains of both proteins is shown. **d**, Genome-wide transcriptional profiles were examined using

M. tuberculosis-specific oligonucleotide microarrays. Gene expression values in *Rv3849*[−] (*R*[−]) and complemented *Rv3849*[−] (*R*⁺) bacteria were divided by expression values in wild-type (WT) bacteria. Results from three independent experiments are shown. Genes with statistically significant changes in expression are shown (see Supplementary Methods for details of statistical analysis). Black, no change; yellow, increased expression; blue, decreased expression relative to wild-type; grey, missing data. Lines to the right of the cluster diagram represent probable operons. **e**, Expression of *Rv3615c* in wild-type and *espR*[−] bacteria in liquid culture (0 h) and 2 h and 4 h following infection of bone-marrow-derived macrophages was measured by quantitative PCR from amplified total *M. tuberculosis* RNA. *Rv3615c* expression was normalized to expression of 16S rRNA. Shown are mean ± s.d. of triplicate measurements from one of 2 experiments. **P* < 0.05 comparing wild type and *espR*[−] at 2 h and at 4 h.

The failure of *espR*[−] mutants to express *Rv3616c–Rv3612c* suggested that the obligatory role of EspR in ESX-1 secretion was to activate transcription of this operon. To test this, we attempted to generate EspR mutants that specifically blocked either its transcriptional activity or its secretion. To this end we made point mutations in the N-terminal DNA-binding region of EspR in residues conserved among EspR homologues in related actinomycetes (Fig. 2c, top). Each of these mutated proteins failed to bind the *Rv3616c–Rv3612c* promoter (Supplementary Fig. 3) and was unable to restore ESAT-6 secretion in *espR*[−] mutant *M. tuberculosis* (Fig. 2d, top). Likewise, since the ESX-1 secretion signal in CFP-10 resides in the carboxy terminus²⁵, we attempted to interrupt EspR secretion by creating a series of C-terminal truncation mutants (Fig. 2c, bottom). Each of these mutants retained full DNA-binding activity, though with variable migration of the bound probe, suggesting alterations in the physical geometry of the protein–DNA complex (Supplementary Fig. 3). Deletion of 10 or more amino acids from the C terminus of EspR led to complete inactivation of ESX-1, while deletion of just the last five amino acids did not prevent ESAT-6 secretion (Fig. 2d, bottom). Each of the variants was expressed at levels equivalent to that of the wild-type protein (data not shown). Gene expression analysis showed that the ability of each mutant to restore ESX-1 secretion correlated with transcriptional control of the EspR regulon (Fig. 2e). Point mutants that failed to bind DNA did not restore wild-type expression of EspR transcriptional targets to *espR*[−] bacteria (Fig. 2e,

lanes 4 and 5). Likewise, EspR lacking five C-terminal amino acids retained partial transcriptional activity, whereas deletion of ten amino acids completely abrogated EspR activity (Fig. 2e, lanes 6 and 7). These data demonstrate that both the DNA-binding and C-terminal domains of EspR are required for transcriptional activity and for ESX-1 function.

Ectopic expression of *Rv3616c–Rv3614c* bypasses EspR

To determine if EspR plays an obligate role in ESX-1 function in addition to its transcriptional activity, we sought to bypass EspR by ectopic expression of the *Rv3616c–Rv3612c* operon in the *espR*[−] mutant. If the sole function of EspR in ESX-1 secretion is to activate this single operon, then ESX-1 secretion should be active under these conditions despite the lack of EspR. Because we had shown previously that *Rv3613c* and *Rv3612c* were not required for ESX-1 (ref. 21), we expressed just *Rv3616c–Rv3614c* under the control of the constitutive GroEL promoter in *espR*[−] bacteria. In this strain, ESX-1 secretion was restored to levels similar to that of wild-type (Fig. 3a, lanes 7 and 8), demonstrating that the sole requisite function of EspR in ESX-1 secretion is the transcriptional activation of *Rv3616c–Rv3614c*. This clearly excludes a role for EspR in subsequent secretory activities, distinguishing it from other co-dependent ESX-1 substrates.

Blocking EspR secretion potentiates ESX-1 gene transcription

As the secretion of EspR is not required for ESX-1 system function, the simplest model is that EspR secretion reduces its activity by

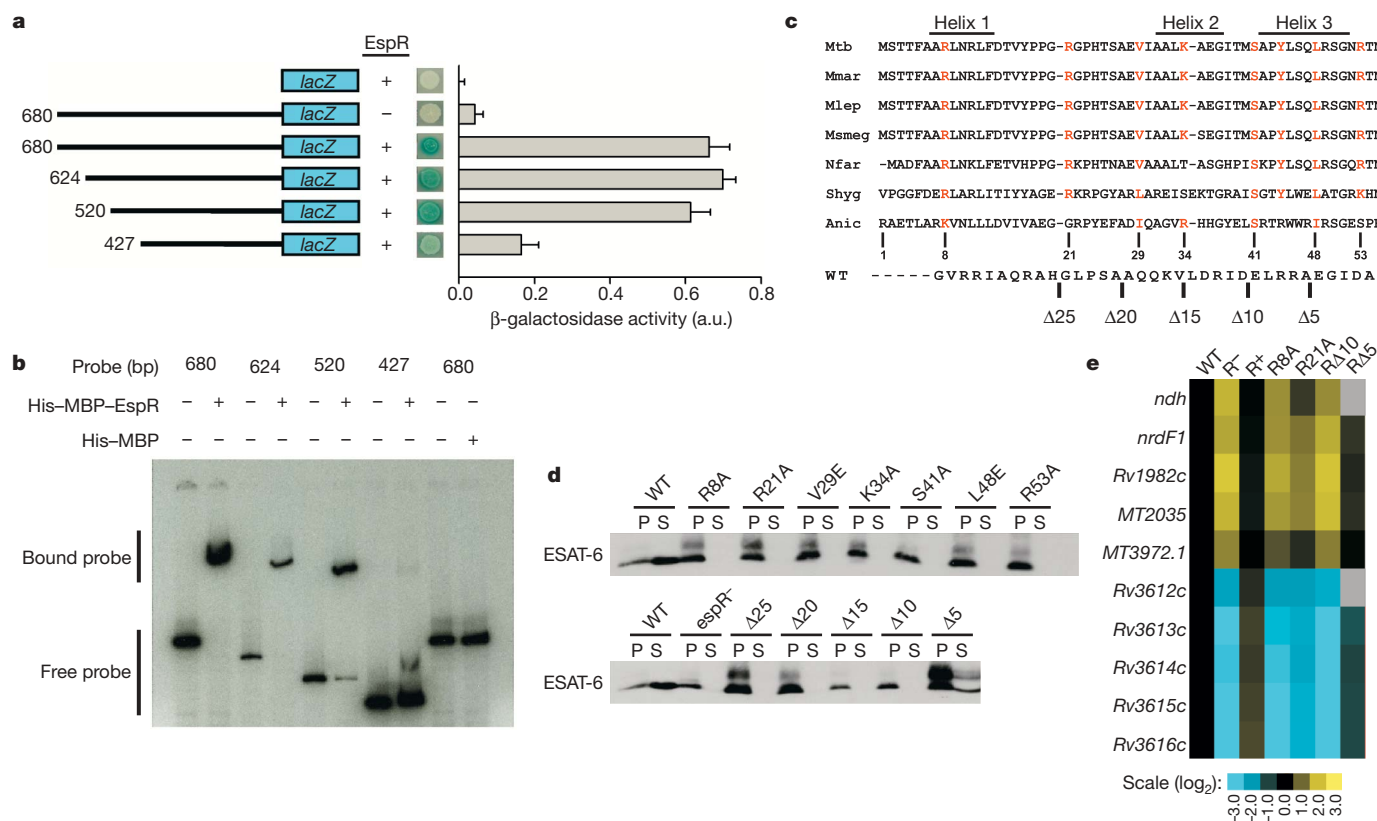


Figure 2 | EspR binds and activates the *Rv3616c–Rv3612c* promoter. **a**, *M. tuberculosis* EspR was expressed in *M. smegmatis* strains carrying *lacZ* reporter constructs with different length *Rv3616c* promoter fragments. Activity was monitored by plating strains on media containing X-gal (5-bromo-4-chloro-3-indolyl-β-D-galactoside) and by quantitative β-galactosidase assay. Shown are mean + s.d. of triplicate measurements from one of 3 experiments. **b**, Binding of EspR to the *Rv3616c* promoter was assayed by EMSA using purified MBP-EspR fusion protein, or MBP alone, and radiolabelled DNA fragments. **c**, N-terminal sequences of EspR homologues of other actinomycetes were aligned and the highly conserved

residues targeted for mutagenesis are highlighted in red (top). A C-terminal deletion series of EspR was also generated (bottom). **d**, Secretion of ESAT-6 from *espR*[−] bacteria expressing mutant forms of EspR was assayed by western blot analysis of cell pellets (P) and culture supernatants (S). **e**, Expression of the EspR regulon in the *espR*[−] mutant (*R*[−]) expressing either wild-type *espR* (*R*⁺) or the indicated mutants was monitored by microarray. Gene expression in each strain was divided by expression level in wild-type (WT) bacteria. Black, no change; yellow, increased expression; blue, decreased expression relative to wild type; grey, missing data.

decreasing the intracellular concentration of the activator. This model predicts that impairment of EspR secretion should result in increased intracellular EspR and resultant increases and decreases in the expression of its positive and negative transcriptional targets, respectively. Although we were unable to isolate a mutant EspR that was specifically impaired for secretion, we serendipitously found that addition of an N-terminal 3×Flag epitope tag to EspR blocked its secretion (Fig. 3b). Using antibodies that recognize EspR, Flag–EspR was detected exclusively in cell pellets, whereas the residual EspR expressed from the mutated chromosomal gene was effectively secreted (Fig. 3b, lanes 3 and 4). Notably, the Flag epitope had little effect on DNA binding to the *Rv3616c–Rv3612c* promoter (Supplementary Fig. 4). Importantly, blocking EspR secretion increased EspR activity: genes activated by EspR were more highly expressed and negatively regulated genes were even further repressed (Fig. 3c). Likewise, titration of intracellular EspR levels using a conditional promoter system led to corresponding increases in *Rv3615c* expression, showing that EspR activity can be modulated by its

intracellular concentration (Fig. 3d, e and Supplementary Fig. 5). This result is consistent with the model that secretion of EspR functions to limit its intracellular activity, functionally coupling transcription and secretion in the ESX-1 system.

Discussion

We have discovered a simple transcriptional feedback mechanism by which secretion of the DNA-binding transcription factor EspR via the ESX-1 pathway negatively regulates this virulence factor secretion system. Secretion of transcriptional regulators is extremely rare in biology, with only three other examples in prokaryotes^{26–29}, none of which bind directly to DNA. In eukaryotes, HIV Tat³⁰ and members of the homeobox family³¹ may be secreted, although the functional importance of their export is unclear. Our data are consistent with the model depicted in Fig. 4, in which the flux of protein secretion through ESX-1 cycles between low- and high-activity states driven by the EspR-regulated activation of *Rv3616c–Rv3612c* and inhibited by secretion of the regulator.

In *M. tuberculosis*, tight control of the ESX-1 system during infection may be achieved through integration of multiple regulators, including both EspR and the two-component regulator PhoP (refs 32, 33). Appropriate activity and timing of ESX-1 secretion is probably critical for survival, as this system is both an essential virulence determinant as well as a major source of potent immunostimulatory antigens. Given the early virulence role of ESX-1 secretion and the antigenicity of ESX-1 substrates to the adaptive immune response, it

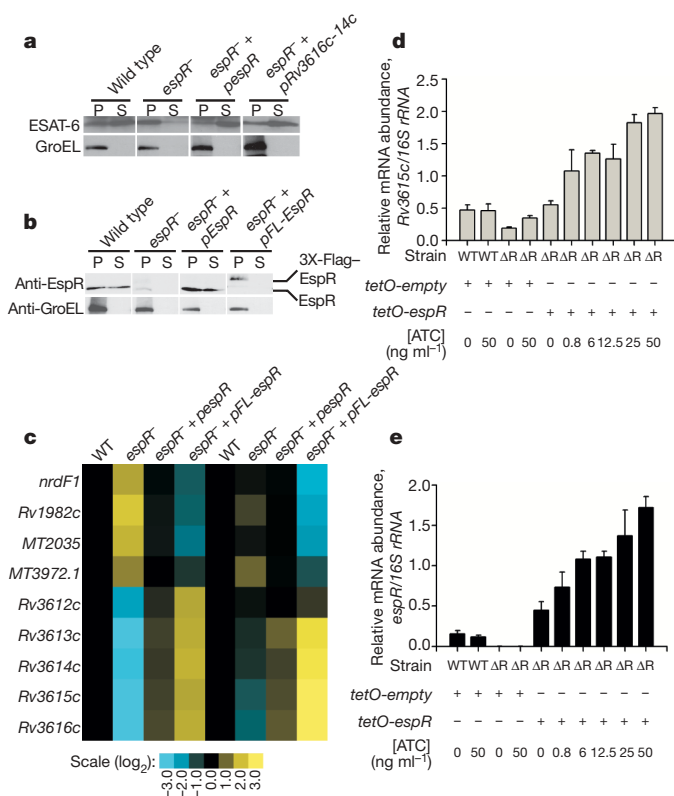


Figure 3 | Examining the role of EspR secretion in ESX-1 function.

a, *Rv3616c–Rv3614c* were expressed in *espR*[−] bacteria under the control of the constitutive *GroEL* promoter. ESAT-6 and GroEL were analysed by western blot in cell pellets (P) and culture supernatants (S) of wild-type bacteria, *espR*[−] bacteria, complemented *espR*[−] bacteria, and *espR*[−] bacteria constitutively expressing *Rv3616c–Rv3614c*. **b**, Cell pellets (P) and culture supernatants (S) of wild-type bacteria and *espR*[−] bacteria expressing N-terminally Flag-tagged EspR were probed by western blot for EspR and GroEL. Arrows indicate endogenous EspR and 3×-Flag–EspR. **c**, Microarray analysis comparing expression of EspR regulon genes in the strains used in **b**. Black, no change; yellow, increased expression; blue, decreased expression relative to wild-type; grey, missing data. Two replicate experiments are shown. **d**, **e**, EspR was placed under the transcriptional control of the Tet-inducible promoter and introduced on an integrating plasmid into wild-type (WT) and Δ espR (Δ R) bacteria. Total RNA was harvested from mid-log phase *M. tuberculosis* grown in the presence of the indicated concentration of anhydrous tetracycline (ATC). *Rv3615c* (**d**) and *espR* (**e**) expression were measured by quantitative real-time PCR and normalized to *16S rRNA* expression. Shown are mean \pm s.d. of triplicate measurements from one representative experiment of three.

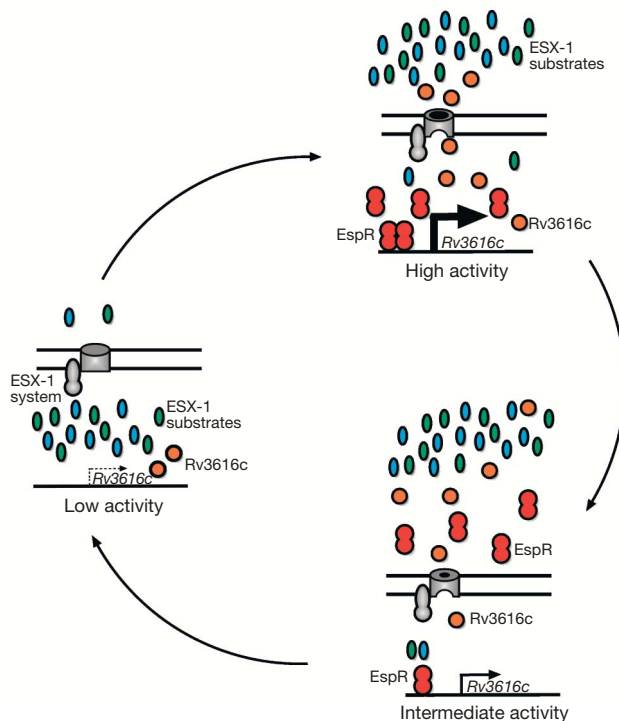


Figure 4 | Model of EspR regulated activation of ESX-1 secretion and negative feedback through export via the ESX-1 pathway. ESX-1 function cycles between low-activity and high-activity states (left and top right, respectively) and is activated by EspR transcriptional activity but limited by EspR secretion. Activation of EspR by unknown stimuli leads to increased transcription (thick black arrow) expression of *Rv3616c* (orange circles) and other genes in the operon (not shown). *Rv3616c* accumulation leads to high activity of ESX-1, perhaps via interactions with membrane-bound ESX-1 secretion components (grey), leading to secretion of substrates, including *Rv3616c*. Secretion of EspR decreases the intracellular concentration of the activator and subsequently lowers *Rv3616c* transcription, depicted here (thin arrow, bottom right) as an intermediate-activity state. Eventually, this negative feedback loop returns the ESX-1 system back to the low-activity state transcription (dashed arrow).

is intriguing to speculate that *M. tuberculosis* may balance the costs and benefits of ESX-1 activity by inducing secretion immediately after phagocytosis infection, through EspR, and then inactivating it as the infection progresses via secretion of the activator. Notably, it has been shown³⁴ that the number of ESAT-6 reactive T cells waned during chronic *M. tuberculosis* infection, a finding consistent with this idea. Understanding how key virulence factors and antigens are regulated during infection deepens our understanding of tuberculosis pathogenesis, and may affect the design of new vaccine strains with enhanced immunogenicity.

METHODS SUMMARY

All strains and plasmids used in this study are described in Supplementary Table 2. Mouse infections, bacterial protein preparation and analysis, RNA preparation, and microarray hybridization were performed as previously described^{1,21,35,36}. EMSAs were performed by incubating purified protein with approximately 10⁵ c.p.m. of radiolabelled probe in the presence of excess unlabelled non-specific DNA before running on a non-denaturing 6% polyacrylamide gel. For qPCR, *M. tuberculosis* cells were isolated from bone-marrow-derived macrophages, total bacterial RNA was isolated and amplified using a Bacterial Amplification Kit (Ambion), reverse transcribed and used as template for PCR in the presence of SYBR Green. Oligonucleotides used for gene amplification are indicated in the Supplementary Methods and expression of individual mRNAs was normalized to expression of *16S rRNA*.

Rv3849 was modelled using the program Modeller (<http://salilab.org/modeller/>) with the crystal structure of SinR as a template (PDB: 1B0N). The Rv3849 model was viewed and manipulated using the program Chimera (<http://www.cgl.ucsf.edu/chimera/>).

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 13 May; accepted 27 June 2008.

- Stanley, S. A., Raghavan, S., Hwang, W. W. & Cox, J. S. Acute infection and macrophage subversion by *Mycobacterium tuberculosis* require a specialized secretion system. *Proc. Natl Acad. Sci. USA* **100**, 13001–13006 (2003).
- Hsu, T. et al. The primary mechanism of attenuation of bacillus Calmette-Guerin is a loss of secreted lytic function required for invasion of lung interstitial tissue. *Proc. Natl Acad. Sci. USA* **100**, 12420–12425 (2003).
- Pathak, S. K. et al. Direct extracellular interaction between the early secreted antigen ESAT-6 of *Mycobacterium tuberculosis* and TLR2 inhibits TLR signaling in macrophages. *Nature Immunol.* **8**, 610–618 (2007).
- Abdallah, A. M. et al. Type VII secretion-mycobacteria show the way. *Nature Rev. Microbiol.* **5**, 883–891 (2007).
- Guinn, K. M. et al. Individual RD1-region genes are required for export of ESAT-6/CFP-10 and for virulence of *Mycobacterium tuberculosis*. *Mol. Microbiol.* **51**, 359–370 (2004).
- MacGurn, J. A. & Cox, J. S. A genetic screen for *Mycobacterium tuberculosis* mutants defective for phagosome maturation arrest identifies components of the ESX-1 secretion system. *Infect. Immun.* **75**, 2668–2678 (2007).
- Volkman, H. E. et al. Tuberculous granuloma formation is enhanced by a mycobacterium virulence determinant. *PLoS Biol.* **2**, e367 (2004).
- Stanley, S. A., Johndrow, J. E., Manzanillo, P. & Cox, J. S. The Type I IFN response to infection with *Mycobacterium tuberculosis* requires ESX-1-mediated secretion and contributes to pathogenesis. *J. Immunol.* **178**, 3143–3152 (2007).
- de Jonge, M. I. et al. ESAT-6 from *Mycobacterium tuberculosis* dissociates from its putative chaperone CFP-10 under acidic conditions and exhibits membrane-lysing activity. *J. Bacteriol.* **189**, 6028–6034 (2007).
- Singh, B., Singh, G., Trajkovic, V. & Sharma, P. Intracellular expression of *Mycobacterium tuberculosis*-specific 10-kDa antigen down-regulates macrophage B7.1 expression and nitric oxide release. *Clin. Exp. Immunol.* **134**, 70–77 (2003).
- Brandt, L., Elhay, M., Rosenkrands, I., Lindblad, E. B. & Andersen, P. ESAT-6 subunit vaccination against *Mycobacterium tuberculosis*. *Infect. Immun.* **68**, 791–795 (2000).
- Colangeli, R. et al. MTS-10, the product of the Rv3874 gene of *Mycobacterium tuberculosis*, elicits tuberculosis-specific, delayed-type hypersensitivity in guinea pigs. *Infect. Immun.* **68**, 990–993 (2000).
- Coler, R. N. et al. Vaccination with the T cell antigen Mtb 8.4 protects against challenge with *Mycobacterium tuberculosis*. *J. Immunol.* **166**, 6227–6235 (2001).
- Dietrich, J. et al. Mucosal administration of Ag85B-ESAT-6 protects against infection with *Mycobacterium tuberculosis* and boosts prior bacillus Calmette-Guerin immunity. *J. Immunol.* **177**, 6353–6360 (2006).
- Kaufmann, S. H. Envisioning future strategies for vaccination against tuberculosis. *Nature Rev. Immunol.* **6**, 699–704 (2006).
- Mahairas, G. G., Sabo, P. J., Hickey, M. J., Singh, D. C. & Stover, C. K. Molecular analysis of genetic differences between *Mycobacterium bovis* BCG and virulent *M. bovis*. *J. Bacteriol.* **178**, 1274–1282 (1996).
- Gao, L. Y. et al. A mycobacterial virulence gene cluster extending RD1 is required for cytotoxicity, bacterial spreading and ESAT-6 secretion. *Mol. Microbiol.* **53**, 1677–1693 (2004).
- Converse, S. E. & Cox, J. S. A protein secretion pathway critical for *Mycobacterium tuberculosis* virulence is conserved and functional in *Mycobacterium smegmatis*. *J. Bacteriol.* **187**, 1238–1245 (2005).
- Brodin, P. et al. Dissection of ESAT-6 system 1 of *Mycobacterium tuberculosis* and impact on immunogenicity and virulence. *Infect. Immun.* **74**, 88–98 (2006).
- Fortune, S. M. et al. Mutually dependent secretion of proteins required for mycobacterial virulence. *Proc. Natl Acad. Sci. USA* **102**, 10676–10681 (2005).
- MacGurn, J. A., Raghavan, S., Stanley, S. A. & Cox, J. S. A non-RD1 gene cluster is required for Snm secretion in *Mycobacterium tuberculosis*. *Mol. Microbiol.* **57**, 1653–1663 (2005).
- Kelley, L. A., MacCallum, R. M. & Sternberg, M. J. Enhanced genome annotation using structural profiles in the program 3D-PSSM. *J. Mol. Biol.* **299**, 499–520 (2000).
- Lewis, R. J., Brannigan, J. A., Smith, I. & Wilkinson, A. J. Crystallisation of the *Bacillus subtilis* sporulation inhibitor SinR, complexed with its antagonist, SinI. *FEBS Lett.* **378**, 98–100 (1996).
- Rohde, K. H., Abramovitch, R. B. & Russell, D. G. *Mycobacterium tuberculosis* invasion of macrophages: Linking bacterial gene expression to environmental cues. *Cell Host Microbe* **2**, 352–364 (2007).
- Champion, P. A., Stanley, S. A., Champion, M. M., Brown, E. J. & Cox, J. S. C-terminal signal sequence promotes virulence factor secretion in *Mycobacterium tuberculosis*. *Science* **313**, 1632–1636 (2006).
- Hughes, K. T., Gillen, K. L., Semon, M. J. & Karlinsey, J. E. Sensing structural intermediates in bacterial flagellar assembly by export of a negative regulator. *Science* **262**, 1277–1280 (1993).
- Rietsch, A., Vallet-Gely, I., Dove, S. L. & Mekalanos, J. J. ExsE, a secreted regulator of type III secretion genes in *Pseudomonas aeruginosa*. *Proc. Natl Acad. Sci. USA* **102**, 8006–8011 (2005).
- Urbanowski, M. L., Lykken, G. L. & Yahr, T. L. A secreted regulatory protein couples transcription to the secretory activity of the *Pseudomonas aeruginosa* type III secretion system. *Proc. Natl Acad. Sci. USA* **102**, 9930–9935 (2005).
- Parsot, C. et al. A secreted anti-activator, OspD1, and its chaperone, Spa15, are involved in the control of transcription by the type III secretion apparatus activity in *Shigella flexneri*. *Mol. Microbiol.* **56**, 1627–1635 (2005).
- Ensolli, B. et al. Release, uptake, and effects of extracellular human immunodeficiency virus type 1 Tat protein on cell growth and viral transactivation. *J. Virol.* **67**, 277–287 (1993).
- Maizel, A. et al. Engrailed homeoprotein secretion is a regulated process. *Development* **129**, 3545–3553 (2002).
- Frigui, W. et al. Control of *M. tuberculosis* ESAT-6 secretion and specific T cell recognition by PhoP. *PLoS Pathog.* **4**, e33 (2008).
- Lee, J. S. et al. Mutation in the transcriptional regulator PhoP contributes to avirulence of *Mycobacterium tuberculosis* H37Ra strain. *Cell Host Microbe* **3**, 97–103 (2008).
- Lazarevic, V., Nolt, D. & Flynn, J. L. Long-term control of *Mycobacterium tuberculosis* infection is mediated by dynamic immune responses. *J. Immunol.* **175**, 1107–1117 (2005).
- Greenstein, A. E. et al. *tuberculosis* Ser/Thr protein kinase D phosphorylates an anti-anti-sigma factor homolog. *PLoS Pathog.* **3**, e49 (2007).
- Schnappiger, D. et al. Transcriptional adaptation of *Mycobacterium tuberculosis* within macrophages: Insights into the phagosomal environment. *J. Exp. Med.* **198**, 693–704 (2003).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We are grateful to S. Stanley for assistance with mouse infections, B. Lim for generation of several purification constructs, and S. Ehrh for providing Tet expression vectors. In addition, we thank H. Madhani for critical reading of the manuscript, and A. DeFranco, A. Sil, E. Brown, R. Locksley, D. Portnoy, D. Monack, P. Nittler, M. Jain, J. MacGurn, and members of the Cox Laboratory for advice and discussions. This work was supported by National Institutes of Health grants (AI63302 and AI51667). J.S.C. acknowledges the support of the Sandler Family Supporting Foundation and the W.M. Keck Foundation.

Author Contributions S.R. performed the secretion assays, virulence studies, protein modeling and β -galactosidase assays. S.R. and P.M. performed the microarray experiments and data analysis. EMSAs were performed by S.R. and K.C. C.D. and P.M. created the Δ espR mutant and performed the Tet-inducible espR experiment. The project was planned and the manuscript written by S.R. and J.S.C.

Author Information All microarray data are available in the Gene Expression Omnibus (GEO) database via accession number GSE11696. Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to J.S.C. (Jeffery.Cox@ucsf.edu).

METHODS

Bacterial strains and plasmids. All *M. tuberculosis* strains used in this study are derived from the wild-type Erdman strain. The Rv3849[−] mutant carries a Tn5370 transposon insertion 13 nucleotides upstream of the Rv3849 initiation codon. See Supplementary Table 2 for list and descriptions of strains and plasmids used in this study.

Mouse infection. C57BL/6 mice were purchased from Charles River Laboratories. All mice were housed and treated humanely, as described in an animal care protocol approved by the UCSF Institutional Animal Care and Use Committee. *M. tuberculosis* cultures were grown to mid-log phase, washed and resuspended in PBS-Tween. Mice were infected intravenously through the tail-vein with 1×10^6 c.f.u. Lungs, spleens and livers were collected, homogenized, and plated for bacterial colonies³⁷. Statistical significance was determined by analysis of variance (ANOVA; $\alpha = 0.05$) and two-tailed *t*-test assuming unequal variance between two groups (wild-type and Rv3849[−]) for each time point.

Macrophage infections. Macrophages were derived from bone marrow progenitors by culturing with macrophage-colony stimulating factor (M-CSF), frozen at -80°C and plated one day before infection³⁸. *M. tuberculosis* cells were grown in Middlebrook 7H9 supplemented with oleic-acid-albumin-dextrose-catalase (OADC), glycerol and 0.05% Tween-80 to mid-log phase. Inocula were washed in PBS, diluted into DMEM supplemented with 10% horse serum, and added to macrophages at a multiplicity of infection of 10. Following phagocytosis for 2 h, macrophages were washed with PBS and incubated in fresh media at 37°C .

Protein preparation and analysis. *M. tuberculosis* strains were grown to mid-log phase in Middlebrook 7H9 media supplemented with 0.05% Tween-80, diluted to $A_{600} = 0.05$ in Sauton's media supplemented with 0.05% Tween-80, and grown to mid-log phase^{37,38}. Cells were washed, diluted to $A_{600} = 0.05$ in Sauton's media supplemented with 0.005% Tween-80 and grown for 5 days in roller bottles. Culture supernatants were collected by centrifugation, filter sterilized, and concentrated. Fifteen micrograms of cell lysates and culture supernatants were separated by SDS-PAGE, and specific proteins were visualized by immunoblotting with antibodies against Rv3849 (mouse polyclonal, ab43676, Abcam), Mpt32 (generated at University of Texas, Southwestern Medical Center, Center for Proteomics Research, Antibody Production Core), ESAT-6 (Hyb 76-8), or GroEL (HAT5) (gifts from P. Andersen).

Protein purification. Coding sequence of wild-type and mutated Rv3849 was cloned into pLIC-HMK (gift from J. Berger). Plasmids were transformed into *E. coli* BL21/DE3/pLysS, and expression was induced for 4 h with 1 mM IPTG. His-tagged MBP-EspR protein was purified using Ni-NTA agarose (Qiagen).

EMSA. Rv3616c promoter sequences were radiolabelled with T4 polynucleotide kinase in the presence of ^{32}P -gamma-ATP, and purified using Sephadex G-50 spin columns. Approximately 10^5 c.p.m. of probe (approximately 50 nM) was incubated with purified protein (20 μM) for 30 min at room temperature in reaction buffer containing 10 mM Tris pH 8.0, 50 mM NaCl, 1 mM EDTA, 5% glycerol, 1 mM DTT, 0.1 mg ml^{−1} salmon sperm DNA (Invitrogen), 25 μg ml^{−1} polydI:dC (Sigma) before running on a non-denaturing 6% polyacrylamide gel. The 680-BP Rv3616c promoter fragment begins at nucleotide position 4057055 in the H37Rv genome sequence (<http://genolist.pasteur.fr/TubercuList/>).

Microarrays. *M. tuberculosis* RNA preparation and microarray hybridization was performed as previously described³⁹. Briefly, total *M. tuberculosis* RNA was purified by bead-beating bacterial pellet in Trizol reagent (Invitrogen), followed by chloroform extraction, isopropanol precipitation, DNase treatment, and cleanup over an RNeasy mini column (Qiagen). 3–5 μg of total RNA was reverse transcribed in the presence of amino-allyl UTP and Cy3- or Cy5-labelled. Competitive hybridizations between Cy5-labelled experimental cDNA and Cy3-labelled reference cDNA were performed for 24 h at 63°C using whole-genome oligonucleotide arrays (Qiagen). Hybridization data were deemed of high quality if they met numerous quality control criteria, including but not limited to minimum spot intensity in each channel, minimum foreground-background differential in each channel, and linear hybridization across a spot in each channel. The Rv3849 regulon was defined as genes that exhibited at least twofold dependence on Rv3849 in at least two out of three experiments. Array results from three independent experiments were analysed using the SAM (Significance Analysis of Microarrays) statistical package⁴⁰ to determine significantly induced or repressed genes. The analysis was performed with a false discovery rate of 0.01.

Quantitative PCR. 1–3 μg of total *M. tuberculosis* mRNA was reverse transcribed and used as template for PCR in the presence of SYBR Green on an Opticon Real-Time PCR Detection System (Bio-Rad Laboratories). Oligonucleotides for amplification of Rv3615c were: oSR603 5'-GAGCGTCTCGGTGTACTG-3' and oSR604 5'-CGTGTCTGTTGAAGTGTGAGC-3', *rrdF1*: oSR601 5'-CAACCTGGGATACCAGCCTG-3' and oSR 602 5'-CATTACGTATGAGCTTCC-3', *16S rRNA*: oAL63 5'-ATGCTACAATCGCCGGTACA-3' and oAL64 5'-GCGTTGCTGATCTGCGATTA-3', *espR*: oPM501 5'-AACCGCCTGTTTCGACACGGTTTAT-3' and oPM502 5'-TTCCTGAGCGTAGCTGTGATAGGT-3'. For each sample, expression of Rv3615c, *rrdF1*, or *espR* was normalized to expression of *16S rRNA*.

For qPCR using RNA isolated from *M. tuberculosis* inside macrophages, bacterial RNA was isolated by first lysing macrophages in guanidine isothiocyanate buffer⁴¹. Intact bacteria were pelleted and washed in PBS-Tween, followed by bead-beating in Trizol, chloroform extraction, and cleanup over an RNeasy mini column. Total bacterial RNA was amplified using a Bacterial Amplification Kit (Ambion), and Rv3615c expression was measured in amplified RNA as above. Statistical significance was determined by ANOVA ($\alpha = 0.05$) and two-tailed *t*-test assuming unequal variance between two groups (wild-type and Rv3849[−]) bacteria.

lacZ reporter assays. Rv3616c promoter fragments were inserted upstream of *lacZ* in the plasmid pYUB76, the promoter-*lacZ* fusions were moved into the integrating plasmid pMV306-Kan and transformed into *M. smegmatis*. *M. tuberculosis* *espR* was cloned downstream of the inducible acetamidase promoter⁴². Owing to sufficient leakiness of the acetamidase promoter, inducer was not used to overexpress *espR*. Doubly-transformed *M. smegmatis* clones were spotted onto plates containing X-gal to monitor *lacZ* expression. For liquid β -galactosidase assays, *M. smegmatis* cultures were grown to $A_{600} = 0.6$ and enzyme activity was measured by the method of ref. 43. β -galactosidase activity in each strain was normalized by subtraction of background activity measured in the strain carrying a plasmid with *lacZ* without an upstream promoter.

Tetracycline inducible expression of EspR. Tetracycline inducible vectors optimized for use in mycobacteria were a gift from S. Ehrh⁴⁴. In *M. tuberculosis*, *ΔespR* bacteria carrying either a empty *tetO* vector or an integrated plasmid with *tetO* driving *ΔespR* expression were grown in liquid culture to $A_{600} = 0.4$ then induced with 0.8–50 ng ml^{−1} of anhydrous tetracycline (ATC, Sigma). At 24 h post-induction, total bacterial RNA was harvested as described for microarrays and analysed by qPCR for *espR* and Rv3615c expression. For *M. smegmatis* experiments, wild-type bacteria containing an integrated copy of the 710-bp Rv3616c promoter-*lacZ* reporter were transformed with an episomal *tetO*-*espR* plasmid. Doubly transformed bacteria were induced for 8 h with 0.8–25 ng ml^{−1} of ATC, and β -galactosidase activity was measured by the method of ref. 43.

Structural modelling. Rv3849 was modelled using the program Modeller (<http://salilab.org/modeller/>) with the crystal structure of SinR as a template (PDB: 1B0N). The Rv3849 model was viewed and manipulated using the program Chimera (<http://www.cgl.ucsf.edu/chimera/>).

- Stanley, S. A., Raghavan, S., Hwang, W. W. & Cox, J. S. Acute infection and macrophage subversion by *Mycobacterium tuberculosis* require a specialized secretion system. *Proc. Natl Acad. Sci. USA* **100**, 13001–13006 (2003).
- MacGurn, J. A., Raghavan, S., Stanley, S. A. & Cox, J. S. A non-RD1 gene cluster is required for Snm secretion in *Mycobacterium tuberculosis*. *Mol. Microbiol.* **57**, 1653–1663 (2005).
- Greenstein, A. E. et al. *tuberculosis* Ser/Thr protein kinase D phosphorylates an anti-anti-sigma factor homolog. *PLoS Pathog.* **3**, e49 (2007).
- Tusher, V. G., Tibshirani, R. & Chu, G. Significance analysis of microarrays applied to the ionizing radiation response. *Proc. Natl Acad. Sci. USA* **98**, 5116–5121 (2001).
- Schnappinger, D. et al. Transcriptional adaptation of *Mycobacterium tuberculosis* within macrophages: Insights into the phagosomal environment. *J. Exp. Med.* **198**, 693–704 (2003).
- Triccas, J. A., Parish, T., Britton, W. J. & Gicquel, B. An inducible expression system permitting the efficient purification of a recombinant antigen from *Mycobacterium smegmatis*. *FEMS Microbiol. Lett.* **167**, 151–156 (1998).
- Timm, J., Lim, E. M. & Gicquel, B. *Escherichia coli*-mycobacteria shuttle vectors for operon and gene fusions to *lacZ*: The pJEM series. *J. Bacteriol.* **176**, 6749–6753 (1994).
- Ehrh, S. et al. Controlling gene expression in mycobacteria with anhydrotetracycline and Tet repressor. *Nucleic Acids Res.* **33**, e21 (2005).

ARTICLES

On the nature of partial agonism in the nicotinic receptor superfamily

Remigijus Lape¹, David Colquhoun¹ & Lucia G. Sivilotti¹

Partial agonists are ligands that bind to receptors but produce only a small maximum response even at concentrations where all receptors are occupied. In the case of ligand-activated ion channels, it has been supposed since 1957 that partial agonists evoke a small response because they are inefficient at eliciting the change of conformation between shut and open states of the channel. We have investigated partial agonists for two members of the nicotinic superfamily—the muscle nicotinic acetylcholine receptor and the glycine receptor—and find that the open–shut reaction is similar for both full and partial agonists, but the response to partial agonists is limited by an earlier conformation change ('flipping') that takes place while the channel is still shut. This has implications for the interpretation of structural studies, and in the future, for the design of partial agonists for therapeutic use.

Agonists are small molecules that bind to a receptor and activate it. The best understood receptors are ligand-gated ion channels. When neurotransmitters bind to their extracellular domain, the resulting change of conformation opens an ion channel, which carries current through the cell membrane, allowing electrical signals to propagate. The natural neurotransmitters, acetylcholine and glycine, are very efficacious agonists: when they are bound, the channel is open for 95–98% of the time^{1,2}. In the 1950s, agonists were discovered that could not produce such a large response even when they saturated the binding sites. These were called partial agonists. Here we find that taurine can hold the glycine receptor channel open for at most about 54% of the time. del Castillo and Katz³ were the first to propose for an ion channel that when the receptor is occupied by a partial agonist the 'gating' equilibrium between open and shut conformations lies towards the shut side. This is equivalent to saying that agonists work because they have a higher affinity for the open state than the shut state^{4,5}, so their binding shifts the equilibrium towards the open state. The more efficacious the agonist, the greater is its selectivity for the open state. These views of partial agonism have persisted, essentially unchanged, for 50 years. However, recent findings suggest another possibility. Φ -value analysis indicates that, after binding, nicotinic receptors move through a number of brief intermediate shut states, a 'conformational wave', before the channel opens^{6–10}. Our own work on glycine receptors suggested that it is possible to detect and measure the properties of an intermediate conformation, which we refer to as 'flip'². The flipped receptor has a higher affinity for the agonist than the resting receptor, so it is a sort of activated, pre-open state. Higher affinity could result from domain closure around the bound agonist, a phenomenon that is clear in structures of extracellular domains of glutamate channels^{11,12}, but less obvious in the nicotinic superfamily^{13,14}, where binding sites are at the interface between subunits.

Here we investigate two partial agonists, taurine for glycine receptors, and tetramethylammonium (TMA) for nicotinic receptors. When our single-channel measurements are interpreted in terms of the flip mechanism, it is found that the open–shut reaction is remarkably similar for full and partial agonists. In both receptors, partial agonism originates from a reduced ability to flip, rather than a reduced ability to open. From the viewpoint of selective affinity, it is a low affinity for the flipped state, relative to the resting state, that

makes an agonist partial, rather than low affinity for the open state, relative to the resting state, as previously supposed. This interpretation places the root of partial agonism early in the chain of events that follow binding. It provides an experimental verification for Φ -value analysis^{6,9,10} as well as a functional counterpart for future structural measurements and it may be exploitable in rational drug design.

Results with the glycine receptor

Figure 1a shows raw single-channel data for three concentrations of a full agonist, glycine, and a partial agonist, taurine, on rat heteromeric $\alpha 1 \beta$ glycine receptors. At the highest concentrations it is obvious that the channel is open for more of the time with glycine than with taurine.

At 1–100 mM taurine, openings occurred in long clusters separated by long shut periods, presumably sojourns in long-lived desensitized states. Below 1 mM, openings were sparse and occurred in short bursts of openings separated by short shuttings (Fig. 1c).

The probability of being open, P_{open} (the fraction of time for which the channel is open; that is, the ratio between total open time per cluster and cluster length), reached a maximum of about 96% for glycine but only 54% for taurine (Fig. 1b). Unlike nicotinic channels, there was no sign of channel block by the agonist itself.

Plotting the distributions of apparent shut times ('apparent' means what is observed, as distorted by our inability to detect short events) shows that for both the partial agonist, taurine, and the full agonist, glycine, the mean lifetime of the brief shuttings is very similar, being between 10 and 14 μs , irrespective of concentration² (Fig. 1d).

To investigate what mechanisms might underlie our observations, we fitted putative mechanisms to the data with our HJCFT method. This method maximizes the likelihood of the entire sequence of openings and shuttings, with exact allowance for missed events. Simulation studies^{2,15} showed that this method is capable of estimating up to 18 rate constants with values as fast as 130,000 s^{-1} .

To estimate all the rate constants it is necessary to fit simultaneously a set of recordings at different agonist concentrations. First the mechanism to be tested must be postulated. We shall concentrate here on the flip mechanism (Fig. 2a) because it provides a realistic description of the wild-type glycine receptor² (see discussion below).

¹Department of Pharmacology, University College London, Medical Sciences Building, Gower Street, London WC1E 6BT, UK.

It was not possible to use taurine concentrations lower than 1 mM, for technical reasons described in Supplementary Information. Thus, it is not surprising that initial fits suggested that there were few monoligated openings in our records. Therefore monoligated flipped and open states were not included in the mechanism in the final fits (Fig. 2a, grey regions).

Figure 2 shows the results of a fit to a set of three records with 3, 10 and 100 mM taurine. It shows that the flip mechanism can, with a

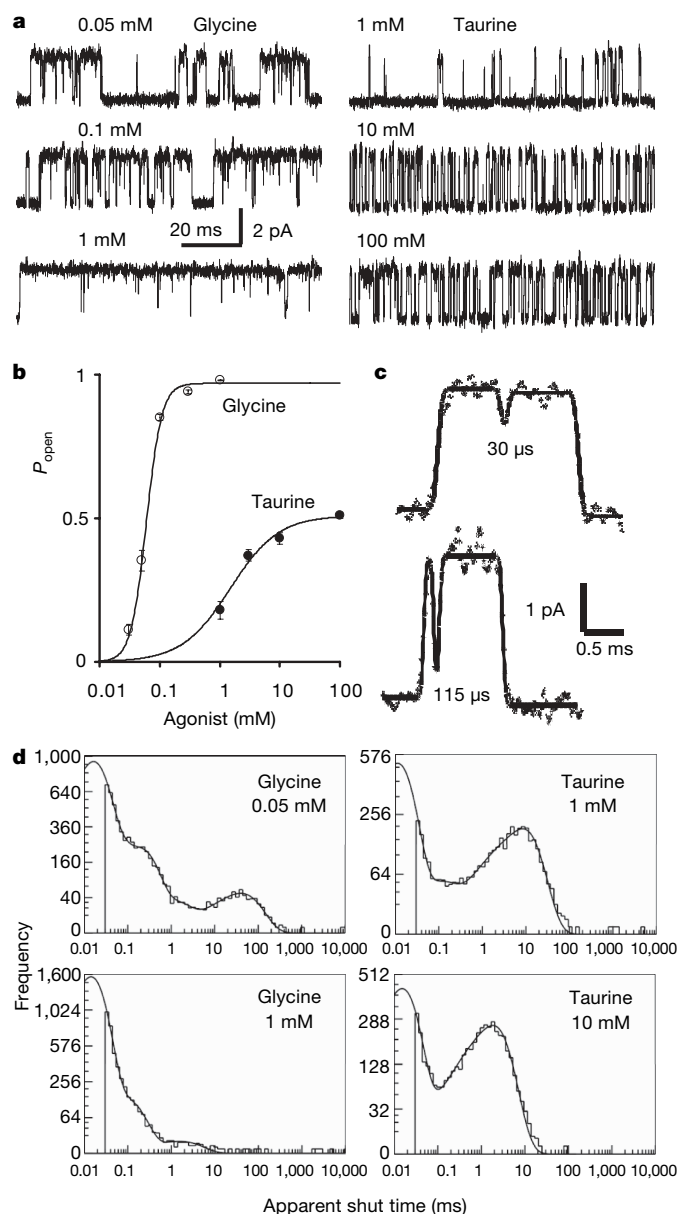


Figure 1 | Glycine channels show short interruptions when activated by either full or partial agonists. **a**, Single-channel currents (opening upwards) produced by a full agonist (glycine) or a partial agonist (taurine). **b**, Open probability increases with agonist concentration, reaching a maximum of 0.54 for taurine and 0.96 for glycine (17–60 clusters, 3–4 patches per point, \pm s.e.m.; Hill equation fits; glycine data from ref. 2). **c**, Short shut times are clearly detectable in channel activations by 1 mM taurine. Points are the digitized record and the continuous line is the best time-course fit obtained during idealization together with the estimated duration of the short gaps. **d**, Short shut times are the predominant component in shut-time distributions for both agonists. With the usual resolution of 30 μ s, a large proportion of such short shuttings are missed, but quite enough are observed to obtain a good estimate of their duration. Dwell-time distributions were fitted here and in Fig. 4b with a mixture of exponentials using EKDIST, so they are descriptive and not mechanism dependent.

single set of rate constants, predict well the observed open and shut times at all three concentrations and the observed P_{open} curve (continuous lines in Fig. 2b, c).

The average rate constants from nine recordings fitted as three independent sets (Table 1 and Supplementary Table 1) show that the opening rate for the fully liganded channel (β_3) is almost the same for the partial agonist, taurine, as it is for glycine. The shutting rate (α_3) is not greatly different from glycine either ($14,500 \text{ s}^{-1}$, compared with $7,000 \text{ s}^{-1}$), so the open–shut equilibrium constants ($E_3 = \beta_3/\alpha_3$) for taurine and glycine are similar (compared with a 180-fold difference in flipping equilibrium constants, Table 1). It has been assumed since 1957 (ref. 3) that partial agonism is a characteristic of the open–shut reaction but that is clearly not the case here. The major difference between taurine and glycine lies at an earlier stage, in the flipping reaction.

For glycine the equilibrium constant for flipping (F_3) is large (27) so most fully liganded receptors will adopt the pre-open flipped conformation. But for taurine this equilibrium constant is only

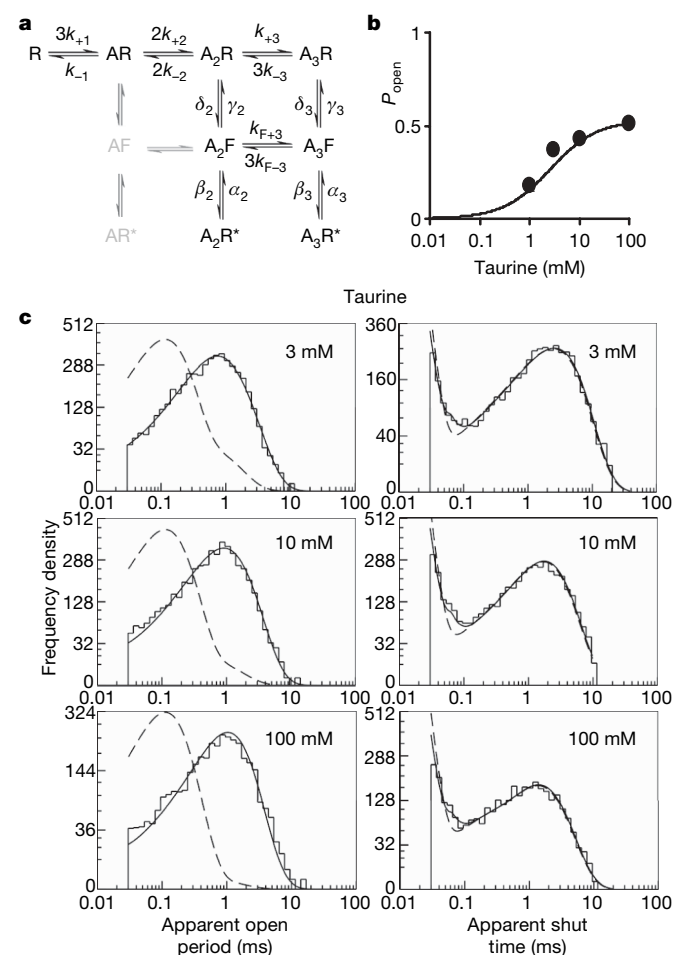


Figure 2 | Global fit of the taurine data with a flip mechanism provides a good description of the observations. **a**, The flip activation mechanism fitted by ref. 2 to $\alpha 1\beta$ channel activations by glycine. ‘A’ represents an agonist molecule; ‘R’ and ‘F’ represent the resting and flipped conformations of the receptor, respectively. The asterisk denotes an open channel. The success of the fit to the idealized records is judged by its ability to describe the experimental observations displayed in various ways¹⁵, as in **b** and **c**. The results of the fit predict accurately the channel open probability (**b**) and apparent open- and shut-time distributions at all concentrations as shown by the continuous lines (fit predictions) superimposed onto the data points and the histograms. The dashed lines in the histograms are the distributions predicted if the records were obtained with perfect resolution (that is, no events were missed; for full data display see Supplementary Fig. 1 and Supplementary Table 1).

Table 1 | Summary of results with full and partial agonists

	Glycine	Taurine	Acetylcholine (−100 mV)	Acetylcholine (+80 mV)	TMA (−80 mV)	TMA (+80 mV)
E_2	-	-	34.4	3.1	28.1	2.8
F_2	-	-	3.8	1.1	0.14	0.060
E_3	20	9.2	-	-	-	-
F_3	27	0.15	-	-	-	-
K (μM)	520	1,040	40.1	43	3,000	2,310
K_F (μM)	8	690	20.7	0.13	340	1,240
α_2 (s ⁻¹)	-	-	2,560	9,550	2,520	9,100
β_2 (s ⁻¹)	-	-	87,700	29,400	70,500	25,000
α_3 (s ⁻¹)	7,000	14,500	-	-	-	-
β_3 (s ⁻¹)	129,000	133,000	-	-	-	-

For estimates of all rate constants, and their errors, and more complete tests of fit, see Supplementary Information. K and K_F are dissociation equilibrium constants for the resting and flipped conformations, respectively.

0.15, so 87% of fully liganded shut receptors will be in the resting conformation, not in the flipped, pre-open conformation. This is summarized in Fig. 3. A schematic representation of the three states (resting, flipped and open) visited by the fully liganded receptor is shown in Fig. 3a. The rates of transition between states are indicated by the size of the arrows (Fig. 3b), for glycine and taurine. The open–shut transition is similar for both agonists. Another way to look at that is through the energy diagram in Fig. 3c, in which the paths for the open–shut transition almost superimpose. The energy diagram shows also that the transition from resting to flipped state is downhill for glycine but uphill for taurine. The overall energy change, from resting to open state, is downhill for glycine, but nearly level for taurine, as expected because the channel is open for about half the time in saturating taurine. In the case of taurine, the transition from resting to flipped is slow, so the mean lifetime of the saturated resting state is much longer than for glycine (Fig. 3b; see also animation in Supplementary Information) and this is why taurine is a partial agonist.

Another way to look at this result is that the effectiveness of an agonist depends largely on its relative affinity for resting and flipped conformations, rather than, as previously supposed, its relative affinity for resting and open conformations. Glycine binds with 65-times greater affinity to the flipped conformation than to the resting state, but taurine binds only slightly (about 1.5-fold) more tightly. Most of this difference arises from the lower affinity of taurine for the flipped state.

Results with the nicotinic receptor

Figure 4a depicts cell-attached activations of human muscle nicotinic receptor elicited by the full agonist acetylcholine (−100 mV transmembrane potential) and by the partial agonist TMA (−80 and +80 mV).

At negative membrane potentials, an acetylcholine-activated channel is practically always open at saturating agonist concentrations (94% of the time), whereas a TMA-bound channel can open for only 78% of the time (Fig. 4c).

TMA recordings showed many brief shuttings, which resemble closely those seen with the full agonist, acetylcholine (mean duration about 13 μs at negative potentials¹) (Fig. 4b). The time constant of the fastest component did not vary noticeably with TMA concentration, but was briefer at negative (12.8 ± 0.46 μs, 13 patches) than at positive (27.9 ± 1.9 μs; 7 patches) membrane potentials. Brief shuttings cannot be attributed entirely to channel block because they were frequent at positive as well as at negative membrane potentials.

Figure 4d shows an important difference between nicotinic and glycine receptors. At negative potentials the amplitude of the openings appears to decrease progressively as agonist concentration increases. This is because in nicotinic receptors the agonist can bind to a second, lower affinity site in the channel pore and produce fast block. Both openings and blockages are very brief and failure to resolve them results in the apparent decrease in channel amplitude (equilibrium constant for TMA block, K_B , 8.9 ± 0.6 mM). At

+80 mV there is little or no block, as expected^{16–18}: channel amplitude is not affected by TMA concentrations as high as 100 mM, the highest on our P_{open} curve. At +80 mV both agonists are less efficacious than at negative potentials: a fully TMA-bound channel opens only for 16% of the time compared with 59% for acetylcholine (Fig. 4a, c and Supplementary Fig. 3).

The original motive for the flip mechanism was to describe in a physically plausible way the apparent strong interaction between binding sites in glycine receptors when conventional mechanisms were fitted². That sort of apparent interaction is much less pronounced for the nicotinic receptor. Nevertheless the flip mechanism fitted nicotinic data well, and it has the virtue of obviating the need to add arbitrary shut states in order to get a good fit^{1,19}.

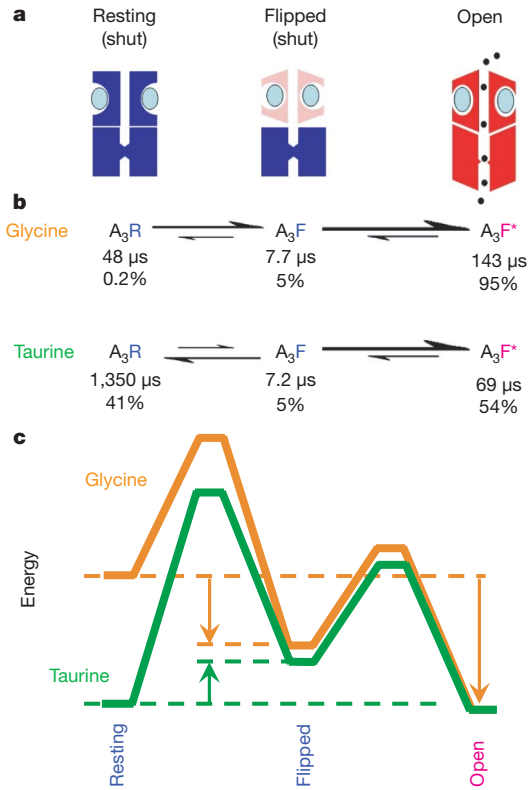


Figure 3 | Activation of the glycine channel by the full agonist glycine and the partial agonist taurine. **a**, The three states of the fully saturated receptor. The resting (shut) channel changes conformation to reach a partially activated state (flipped), still shut, but with increased affinity for the agonist (blue ellipses). It is from this flipped state that the channel can open (red, right). **b**, The reaction rates for channel activation by glycine and taurine differ largely in the first step, the flipping. Relative rates are indicated by the size of arrows. The diagram shows also the mean lifetime of each state (μs) and the proportion of time (percentage) spent by a bound channel in each of the states at equilibrium. A channel occupied by glycine doesn't stay long in the resting state, but quickly flips. When taurine is bound, the channel takes a long time to flip, so it spends much more time in the closed resting state. Once the channel has reached the flipped intermediate state, it opens quickly, regardless of which agonist is bound. The probability of opening, rather than unflipping, is over 96% for both agonists, so a burst of many openings results. **c**, An energy diagram for the three states of the saturated receptor, for full agonist (glycine, orange) and partial agonist (taurine, green). Again, the difference lies largely in the resting–flipped transition. As indicated by arrows, the transition from resting to flipped is downhill for glycine, but uphill for taurine. The overall transition from resting to open is very downhill for glycine, but almost level for taurine, as expected from the maximum response of about 54% open. The calculations used a frequency factor²⁷ of 10⁷ s⁻¹ and the lines are shifted vertically so that they meet at the open state. (See also Supplementary Movie.)

Figure 5 shows part of the results of fitting TMA records with the mechanism in Fig. 5a (see also Supplementary Figs 4 and 5 and Table 1). At -80 mV the largest TMA concentration used for fits was 30 mM because channel block became too pronounced at higher concentrations (Fig. 4d). Again the distributions of apparent dwell times and the P_{open} curve are predicted well by the results of the fit (Fig. 5b, c).

The fully liganded opening rate (β_2) at -80 mV is about $71,000 \text{ s}^{-1}$, very similar to the value for acetylcholine ($88,000 \text{ s}^{-1}$).

The shutting rates also are almost identical, with $\alpha_2 = 2,560 \text{ s}^{-1}$ for acetylcholine and $2,520 \text{ s}^{-1}$ for TMA. Consequently, the fully liganded gating constant, E_2 , is similar for acetylcholine and TMA. Again, what differs between full and partial agonist is the flipping step. For acetylcholine the flipping equilibrium constant, F_2 , is 3.8 but for TMA it is only 0.14, so TMA has a far smaller ability to elicit the pre-open flipped state than acetylcholine, and that is why it is a partial agonist.

It is important to rule out the possibility that these results could be influenced by channel block because at the highest TMA concentration

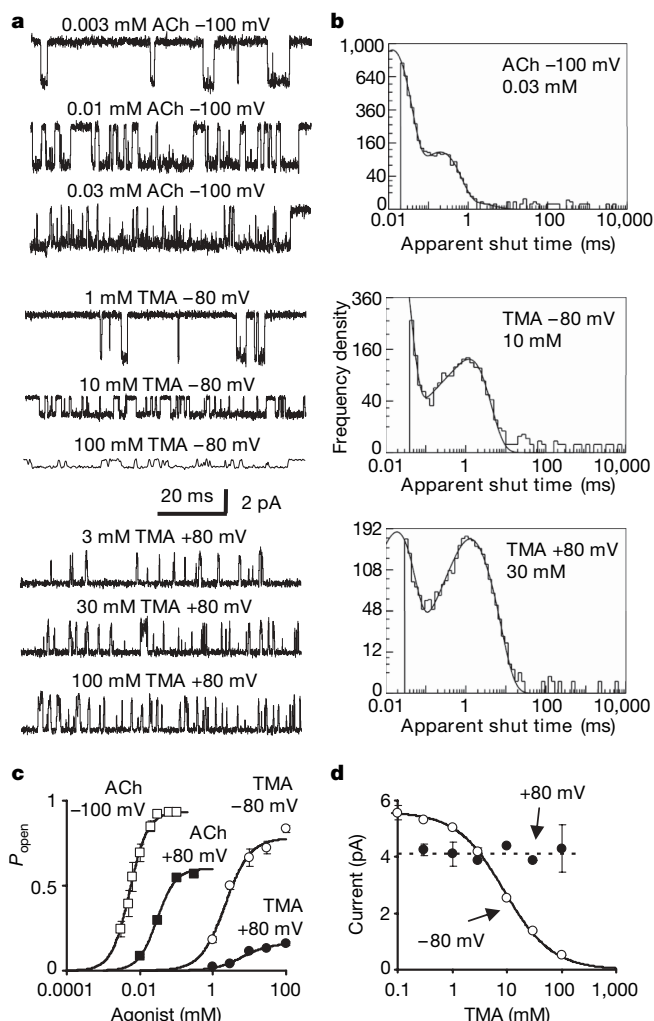


Figure 4 | Short interruptions in openings of muscle nicotinic channels occur with both partial and full agonists and cannot be attributed entirely to channel block. **a**, Single channel currents produced by a full agonist (acetylcholine) and a partial agonist (TMA, openings downwards in the first six rows). ACh, acetylcholine. **b**, Short shut times are a major feature of shut-time distributions irrespective of agonist, voltage and concentration (see Supplementary Figs 2 and 4). Some of these brief shuttings could result from fast open-channel block at negative potentials, but not at positive potential, when open-channel block is essentially absent. **c**, Open probability reaches a maximum of $0.94 (\pm 0.004)$ for acetylcholine at -100 mV but is lower (0.59 ± 0.01) for acetylcholine at $+80$ mV and for TMA at -80 or $+80$ mV (0.78 ± 0.05 or 0.16 ± 0.02 , respectively) (2–4 patches per point, \pm s.e.m.; Hill equation fits). These fits give also: for acetylcholine -100 mV, concentration for 50% of maximum response (EC_{50}) = $5.3 \pm 0.2 \mu\text{M}$ and Hill coefficient (n_H) = 1.7 ± 0.1 ; for acetylcholine $+80$ mV, EC_{50} = $29 \pm 2 \mu\text{M}$ and n_H = 1.7 ± 0.1 ; for TMA -80 mV, EC_{50} = $2.2 \pm 0.5 \text{ mM}$ and n_H = 1.3 ± 0.4 ; and for TMA $+80$ mV, EC_{50} = $6.4 \pm 1.9 \text{ mM}$ and n_H = 1.2 ± 0.4 . **d**, Because of open-channel block, the amplitude of single-channel currents declines with increasing TMA concentrations at negative potentials (open circles), but not at positive transmembrane potentials (filled circles, 2–3 patches per point, \pm s.e.m.).

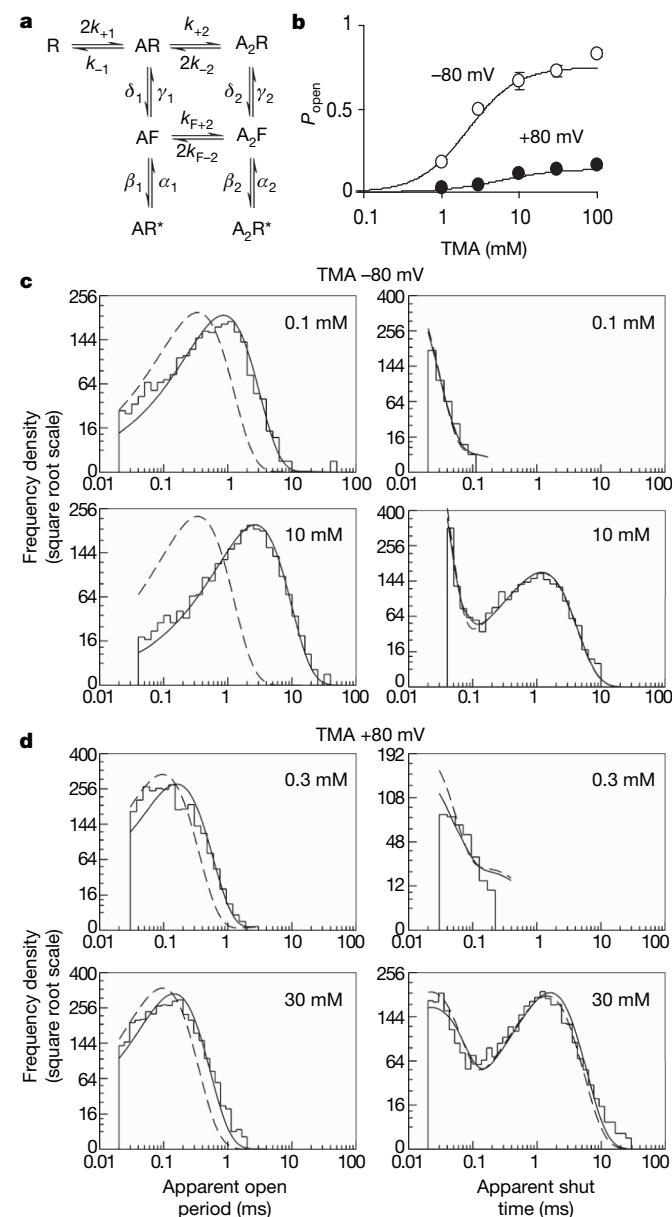


Figure 5 | The flip mechanism describes well the activation of acetylcholine receptors by TMA. **a**, Schematic representation of the flip mechanism used here: this has two agonist binding sites. As in Fig. 2, the experimental data are displayed as channel open probability values (\pm s.e.m.) in **b** or dwell-time distributions in **c**, **d**. The results of fitting the idealized records are used to calculate the predicted P_{open} —concentration curves and distributions (continuous lines). Both the open probability values and the dwell time distributions are well described by the fitted rate constants for TMA, at both -80 mV (**c**) and $+80$ mV (**d**). The dashed lines in the histograms are the distributions predicted if the records were obtained with perfect resolution (that is, no events were missed). The full data display for the fits for acetylcholine and TMA is in Supplementary Figs 2, 4 and 5 and Supplementary Table 2.

fitted (10 mM), channels are blocked for about half the time at -80 mV. The experiments were therefore repeated at positive membrane potential, $+80$ mV, where block is essentially absent (Fig. 4d). Short shuttings are still present at positive membrane potentials, so they cannot all be caused by brief channel blockages. The open–shut equilibrium constant is reduced about tenfold at $+80$ mV (Table 1), so acetylcholine itself becomes a partial agonist¹⁸ (Fig. 4c and Supplementary Fig. 3). The maximum P_{open} for acetylcholine is reduced from 94% at -100 mV to about 60% at $+80$ mV and the maximum P_{open} for TMA is reduced from 78% to 16%. So TMA is still much less efficacious than acetylcholine when the complication of block is removed, and for exactly the same reason. The open–shut equilibrium constant for TMA at $+80$ mV ($E_2 = 2.8$) is almost the same as that for acetylcholine ($E_2 = 3.1$), but again the flipping equilibrium constant for TMA, F_2 , is very small, only 0.06, almost 20-fold less than for acetylcholine, so this remains the reason why TMA is a partial agonist.

Discussion

Our analysis of single-channel data suggests that in the nicotinic superfamily partial agonism is not, as has been supposed for 50 years, a property of the open–shut transition, but arises in the earlier conformation change from the resting state to an intermediate pre-open (or ‘activated’) conformation dubbed the flip state. The same conclusion was reached for two members of this superfamily: the nicotinic acetylcholine receptor and the heteromeric glycine receptor. Figure 3 represents schematically the states visited by the agonist-bound channel during activation and summarizes our quantitative findings for the glycine receptor.

The conclusion that the final step in activation, channel opening, is much the same for full and partial agonists comes from the observation that both partial and full agonists produce similar brief shuttings (7–8 μ s on average for both glycine and taurine). It has been supposed since 1981 (refs 20, 21) that these brief closures are mostly oscillations between the open state and the immediately preceding fully liganded shut state (here the flipped intermediate), and therefore reflect the properties of the open–shut reaction. The similarity in the open–shut transition for the two agonists is obvious from the fact that the energy profile for this step is almost identical for glycine and taurine (Fig. 3c).

It has been known for a long time²² that the conductance of single channels is the same whichever agonist activates them. It is intriguing that we now find that not only the conductance but also the opening and shutting rate of channels do not vary greatly between full and partial agonists in the nicotinic superfamily.

The maximum P_{open} response that can be obtained has always been assumed to depend only on the equilibrium constant for the open–shut step (E), being $E/(1 + E)$ (ref. 23). In the context of the flip mechanism, the maximum response depends not only on E but also on the equilibrium constant for flipping, F . The maximum P_{open} becomes $E_{\text{eff}}/(1 + E_{\text{eff}})$, where $E_{\text{eff}} = EF/(1 + F)$. Here this effective gating constant, E_{eff} , is about 19 for glycine but only 1.2 for taurine, largely because F is lower for the partial agonist. For taurine, the early conformational step is uphill energetically (green arrow in Fig. 3c), and it is much slower than it is for glycine (illustrated by the size of the arrows in Fig. 3b; Supplementary Table 1). The equilibrium between resting and flipped favours the resting state, in which taurine-bound channels spend much more time (on average 1,350 μ s) than glycine-bound channels (48 μ s).

Our interpretation places the origin of partial agonism at an earlier stage in the activation process than before. This is entirely consistent with the conclusion of a previous study⁶, in which the analysis of linear free-energy relationships implied that the differences between nicotinic agonists lie early in the activation process, quite close to the transmitter binding site. Indeed calculation of the Φ value for the overall resting-to-open transition gives a value close to 1, as that study found⁶ (Supplementary Information). Our result also explains

the observation that the conformation change close to the membrane, in the M2–M3 region, appears to be similar for both glycine and taurine, as judged by the reaction rate with a cysteine-reactive compound²⁴.

We may speculate that flipping corresponds to the domain closure around the bound agonist seen in the extracellular domain of glutamate channels and, to a lesser extent, in the acetylcholine binding protein^{11–14}. This structural work cannot be compared directly with our results because it all comes from non-functional receptors, does not tell us how fast domain closure is, how it changes agonist affinity in the intact channel or how it relates to channel opening.

Our single-channel work allowed us to obtain a functional characterization of a partially activated intermediate state, to measure exit and entry rates into this state and its average lifetime with different agonists in two nicotinic superfamily channels. This finding will constrain and inform computational simulations of the dynamics of channel activation and influence the interpretation of structural studies. In addition to that, our findings may also help reinterpret earlier mutation studies, as the greater detail of our model now shows that what was called ‘affinity for the resting state’ may reflect also the higher affinity for a partially activated intermediate. This could perhaps account for the awkward observation that some mutations that are nowhere near the binding site nevertheless appear to affect largely the agonist binding to the resting state²⁵. Another implication testable by structural data is that partial and full agonists should differ in their interaction with the portion of the binding site that moves during domain closure.

METHODS SUMMARY

Cell-attached single-channel currents were recorded from HEK293 cells expressing²⁶ human adult muscle nicotinic receptors or rat $\alpha 1\beta$ glycine receptors.

Rate constant values were estimated by direct model fitting to sets of records idealized by time-course fitting (HJCFT and SCAN programs, respectively; <http://www.ucl.ac.uk/Pharmacology/dc.html>).

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 24 April; accepted 5 June 2008.

Published online 16 July 2008.

1. Hatton, C. J., Shelley, C., Brydson, M., Beeson, D. & Colquhoun, D. Properties of the human muscle nicotinic receptor, and of the slow-channel myasthenic syndrome mutant ϵ L221F, inferred from maximum likelihood fits. *J. Physiol.* **547**, 729–760 (2003).
2. Burzomato, V., Beato, M., Groot-Kormelink, P. J., Colquhoun, D. & Sivilotti, L. G. Single-channel behavior of heteromeric $\alpha 1\beta$ glycine receptors: An attempt to detect a conformational change before the channel opens. *J. Neurosci.* **24**, 10924–10940 (2004).
3. del Castillo, J. & Katz, B. Interaction at end-plate receptors between different choline derivatives. *Proc. R. Soc. Lond. B* **146**, 369–381 (1957).
4. Wyman, J. & Allen, D. W. The problem of the heme interactions in hemoglobin and the basis of the Bohr effect. *J. Polym. Sci. C* **7**, 499–518 (1951).
5. Grosman, C., Zhou, M. & Auerbach, A. Mapping the conformational wave of acetylcholine receptor channel gating. *Nature* **403**, 773–776 (2000).
6. Chakrapani, S., Bailey, T. D. & Auerbach, A. Gating dynamics of the acetylcholine receptor extracellular domain. *J. Gen. Physiol.* **123**, 341–356 (2004).
7. Zhou, Y., Pearson, J. E. & Auerbach, A. ϕ -value analysis of a linear, sequential reaction mechanism: theory and application to ion channel gating. *Biophys. J.* **89**, 3680–3685 (2005).
8. Auerbach, A. Gating of acetylcholine receptor channels: brownian motion across a broad transition state. *Proc. Natl Acad. Sci. USA* **102**, 1408–1412 (2005).
9. Purohit, P., Mitra, A. & Auerbach, A. A stepwise mechanism for acetylcholine receptor channel gating. *Nature* **446**, 930–933 (2007).
10. Armstrong, N., Sun, Y., Chen, G. Q. & Gouaux, E. Structure of a glutamate-receptor ligand-binding core in complex with kainate. *Nature* **395**, 913–917 (1998).
11. Jin, R., Banke, T. G., Mayer, M. L., Traynelis, S. F. & Gouaux, E. Structural basis for partial agonist action at ionotropic glutamate receptors. *Nature Neurosci.* **6**, 803–810 (2003).
12. Celie, P. H. et al. Nicotine and carbamylcholine binding to nicotinic acetylcholine receptors as studied in AChBP crystal structures. *Neuron* **41**, 907–914 (2004).
13. Hansen, S. B. et al. Structures of aplysia AChBP complexes with nicotinic agonists and antagonists reveal distinctive binding interfaces and conformations. *EMBO J.* **24**, 3635–3646 (2005).

15. Colquhoun, D., Hatton, C. J. & Hawkes, A. G. The quality of maximum likelihood estimates of ion channel rate constants. *J. Physiol.* **547**, 699–728 (2003).
16. Adams, P. R. Voltage jump analysis of procaine action at frog end-plate. *J. Physiol.* **268**, 291–318 (1977).
17. Neher, E. & Steinbach, J. H. Local anaesthetics transiently block currents through single acetylcholine-receptor channels. *J. Physiol.* **277**, 153–176 (1978).
18. Colquhoun, D. & Ogden, D. C. Activation of ion channels in the frog end-plate by high concentrations of acetylcholine. *J. Physiol.* **395**, 131–159 (1988).
19. Salamone, F. N., Zhou, M. & Auerbach, A. A re-examination of adult mouse nicotinic acetylcholine receptor channel activation kinetics. *J. Physiol.* **516**, 315–330 (1999).
20. Colquhoun, D. & Sakmann, B. Fluctuations in the microsecond time range of the current through single acetylcholine receptor ion channels. *Nature* **294**, 464–466 (1981).
21. Colquhoun, D. & Sakmann, B. Fast events in single-channel currents activated by acetylcholine and its analogues at the frog muscle end-plate. *J. Physiol. (Lond.)* **369**, 501–557 (1985).
22. Gardner, P., Ogden, D. C. & Colquhoun, D. Conductances of single ion channels opened by nicotinic agonists are indistinguishable. *Nature* **309**, 160–162 (1984).
23. Colquhoun, D. Binding, gating, affinity and efficacy. The interpretation of structure-activity relationships for agonists and of the effects of mutating receptors. *Br. J. Pharmacol.* **125**, 923–948 (1998).
24. Han, N. L., Clements, J. D. & Lynch, J. W. Comparison of taurine- and glycine-induced conformational changes in the M2–M3 domain of the glycine receptor. *J. Biol. Chem.* **279**, 19559–19565 (2004).
25. Colquhoun, D., Unwin, N., Shelley, C., Hatton, C. & Sivilotti, L. G. *Drug Discovery and Drug Development* (ed. Abrahams, D.) 357–405 (John Wiley, New York, 2003).
26. Groot-Kormelink, P. J., Beato, M., Finotti, C., Harvey, R. J. & Sivilotti, L. G. Achieving optimal expression for single channel recording: a plasmid ratio approach to the expression of $\alpha 1$ glycine receptors in HEK293 cells. *J. Neurosci. Methods* **113**, 207–214 (2002).
27. Andersen, O. S. Editorial: Graphic representation of the results of kinetic analyses. *J. Gen. Physiol.* **114**, 589–590 (1999).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements This work was supported by grants from the MRC (Programme grant G0400869) and the Wellcome Trust (Project grant 074491) to L.G.S. and D.C. We are grateful to F. Abogadie for molecular biology, to I. Vais for programming help, and to D. Jane for purification of taurine.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to D.C. (d.colquhoun@ucl.ac.uk).

METHODS

Single-channel electrophysiology. Cell-attached single-channel currents were recorded at 19–21 °C from HEK293 cells transiently transfected by a Ca^{2+} -phosphate co-precipitation method²⁶ with pcDNA3.1 plasmids coding for human muscle acetylcholine nicotinic receptor subunits ($\alpha 1$, β , δ , ϵ , ratio 2:1:1:1) or rat heteromeric glycine receptor subunits ($\alpha 1$, β , ratio 1:1 or 1:4), together with plasmid coding for eGFP. The pipette solution was freshly prepared by diluting agonist stocks with extracellular solution which contained (in mM): 5.4 NaCl, 142 KCl, 1.8 CaCl_2 , 1.7 MgCl_2 and 10 HEPES (pH adjusted to 7.4 with KOH) for acetylcholine receptor experiments or, for glycine receptor experiments, 102.7 NaCl, 20 Na gluconate, 4.7 KCl, 2 CaCl_2 , 1.2 MgCl_2 , 10 HEPES, 14 glucose, 15 sucrose, 20 TEA-Cl (pH adjusted to 7.4 with NaOH). Osmolarity was 320 mOsm for all solutions.

Taurine (Fluka) was found by an HPLC assay to be contaminated by glycine (3 parts in 100,000, molar ratios). It was, therefore, purified before use (by D. Jane). Taurine was passed through an ion exchange resin column, followed by crystallization from the water eluate. This purification was found to be essential to avoid glycine-induced openings in recordings made at high taurine concentrations.

Patch pipettes were made from thick-walled borosilicate glass (GC150F, Harvard Instruments) and coated near the tip with Sylgard 184 (Dow Corning). Electrode resistance was in the range 8–15 M Ω after fire-polishing. Single-channel currents were recorded with an Axopatch 200B amplifier (Axon Instruments), filtered at 10 kHz (sampling rate 100 kHz) and stored on the PC hard drive.

Pipette potential was held at +100 mV for the glycine experiments and at +100, –80 or +80 mV for the nicotinic experiments.

Analysis. For off-line analysis, data were filtered at 3–8 kHz and sampled at 33–100 kHz. Records were idealized by time-course fitting (SCAN program; <http://www.ucl.ac.uk/Pharmacology/dc.html>), and the appropriate resolution (20–100 μs) was imposed before fitting dwell-time distributions with mixtures of exponential probability densities (EKDIST program). Rate constant values were estimated by direct model fitting to the idealized records (HJCFIT program). This was performed on sets of experiments, each set containing three (four for acetylcholine) agonist concentrations spanning the dose–response curve from the minimum concentration that elicited clusters of openings (that could be attributed to a single molecule) to the maximum P_{open} . Each experiment contained 10,000–20,000 transitions. Fits were repeated, systematically changing initial guesses to ensure that the convergence was to a global maximum of the likelihood surface.

ARTICLES

A mechanism for asymmetric segregation of age during yeast budding

Zhanna Shcheprova¹, Sandro Baldi¹, Stephanie Buvelot Frei¹, Gaston Gonnet² & Yves Barral¹

Ageing and the mortality that ensues are sustainable for the species only if age is reset in newborns. In budding yeast, buds are made young whereas ageing factors, such as carbonylated proteins and DNA circles, remain confined to the ageing mother cell. The mechanisms of this confinement and their relevance are poorly understood. Here we show that a septin-dependent, lateral diffusion barrier forms in the nuclear envelope and limits the translocation of pre-existing nuclear pores into the bud. The retention of DNA circles within the mother cell depends on the presence of the diffusion barrier and on the anchorage of the circles to pores mediated by the nuclear basket. In accordance with the diffusion barrier ensuring the asymmetric segregation of nuclear age-determinants, the barrier mutant *bud6Δ* fails to properly reset age in buds. Our data involve septin-dependent diffusion barriers in the confinement of ageing factors to one daughter cell during asymmetric cell division.

The propagation of metazoans relies on the separation of the soma, which ages and is mortal, from the germinal lineage, which survives through generations. How the germline is protected from ageing is not understood. Remarkably, unicellular organisms as distinct as *Escherichia coli*¹ and budding yeast² are also subjected to ageing, and provide a simple system to study how cells reset age.

Baker's yeast cells divide asymmetrically through budding of a daughter cell from the surface of its mother³. Mother and daughter cells are profoundly different. The mother devotes biosynthesis to the generation of daughters, and ages^{4,5}: it loses fitness with time, its division rate decreases and finally it dies^{6,7}. In contrast, daughters dedicate biosynthesis to their own growth until they reach the critical size and turn into mothers themselves⁸. Fairly irrespectively of the age of their mother, the average lifespan of buds remains constant⁹, indicating that somehow their ages are reset.

Several processes contribute to yeast ageing^{6,7}. One is the accumulation of extrachromosomal ribosomal DNA circles (ERCs) in mother nuclei. These episodes sporadically pop out of chromosomal ribosomal DNA through homologous recombination, replicate during the synthesis phase and segregate asymmetrically to the mother cell during mitosis¹⁰. Why their accumulation contributes to the ageing of the mother is unclear. However, any replicating episome lacking segregation sequences such as a centromere (called non-centromeric episomes) segregates to the mother cell¹¹, and accelerates ageing¹². These observations indicate that the yeast nucleus divides somewhat asymmetrically, and that this asymmetry protects daughters from inheriting ageing factors such as non-centromeric episomes.

A diffusion barrier in the outer membrane of the anaphase nucleus

To investigate the mechanisms underlying the asymmetric division of yeast nuclei, we first used fluorescence loss in photobleaching (FLIP) to investigate the dynamics of the different nuclear compartments during anaphase. Green fluorescent protein (GFP)-tagged reporters were repeatedly photobleached in a small area, and fluorescence was monitored throughout the nucleus. Fluorescence decay

reported on molecular exchange between the location monitored and the photobleached area¹³. The outer-membrane components Sec61 (ref. 14) and Nsg1 (ref. 15), the nucleoporin Nup49 (ref. 16) and the inner-membrane protein Prm3 (ref. 17) reported on the dynamics of the nuclear envelope. Sec61 is a subunit of the translocon, Nsg1 is an integral membrane protein and Prm3 is involved in nuclear fusion during mating. GFP fused to the endoplasmic reticulum (ER)-retrieval sequence histidine-aspartate-glutamate-leucine¹⁸ (GFP-HDEL) reported on the perinuclear lumen. GFP fused to a nuclear localization sequence (NLS-GFP)¹⁹ reported on the nucleoplasm.

Prior to anaphase, photobleaching of Nup49-GFP in a small area of the nuclear envelope led to rapid fluorescence decay over the entire nucleus (Supplementary Fig. 2a). Thus, unlike nuclear pore complexes (NPCs) of animal cells²⁰, and as already reported^{21,22}, the yeast NPCs moved freely throughout the nuclear envelope. However, as soon as a nuclear lobe penetrated into the bud (initial translocation; see Supplementary Fig. 1 for the different stages of nuclear division), photobleaching of a nuclear spot in the mother caused rapid extinction of the entire mother lobe but little fluorescence loss in the bud fraction of the nucleus (Fig. 1a, b, Supplementary Movie 1, Supplementary Fig. 2b–d). This lack of fluorescence loss in the bud was observed throughout early, mid and late anaphase (Fig. 1a, Supplementary Fig. 2b–d). Similarly, photobleaching in the bud fraction of early-anaphase nuclei caused rapid fluorescence loss exclusively in the bud (Supplementary Fig. 2e). Thus, upon penetration of the nucleus into the bud, the nuclear envelope became compartmentalized with respect to NPC movement, with a sharp diffusion boundary at the bud neck. This compartmentalization was also observed for the other outer-membrane markers Nsg1-GFP and Sec61-GFP (Fig. 1c, Supplementary Fig. 3a, Supplementary Movies 2, 3). In contrast, exchange between mother and bud remained rapid for the nucleoplasm (NLS-GFP; Supplementary Fig. 3b, Supplementary Movie 4), the inner membrane (Prm3-GFP; Supplementary Fig. 3c, Supplementary Movie 5) and the ER lumen (GFP-HDEL; Supplementary Fig. 3d, Supplementary Movie 6).

Quantification indicated that the decay of Nup49-GFP fluorescence was 24.7 ± 10.7 ($n = 13$) times faster in the bleached

¹Institute of Biochemistry, Biology Department, ETH Zurich, Schafmattstrasse 18, 8093 Zurich, Switzerland. ²Institute of Computational Science, Department of Computer Science, ETH Zurich, Universitätsstrasse 6, 8092 Zürich, Switzerland.

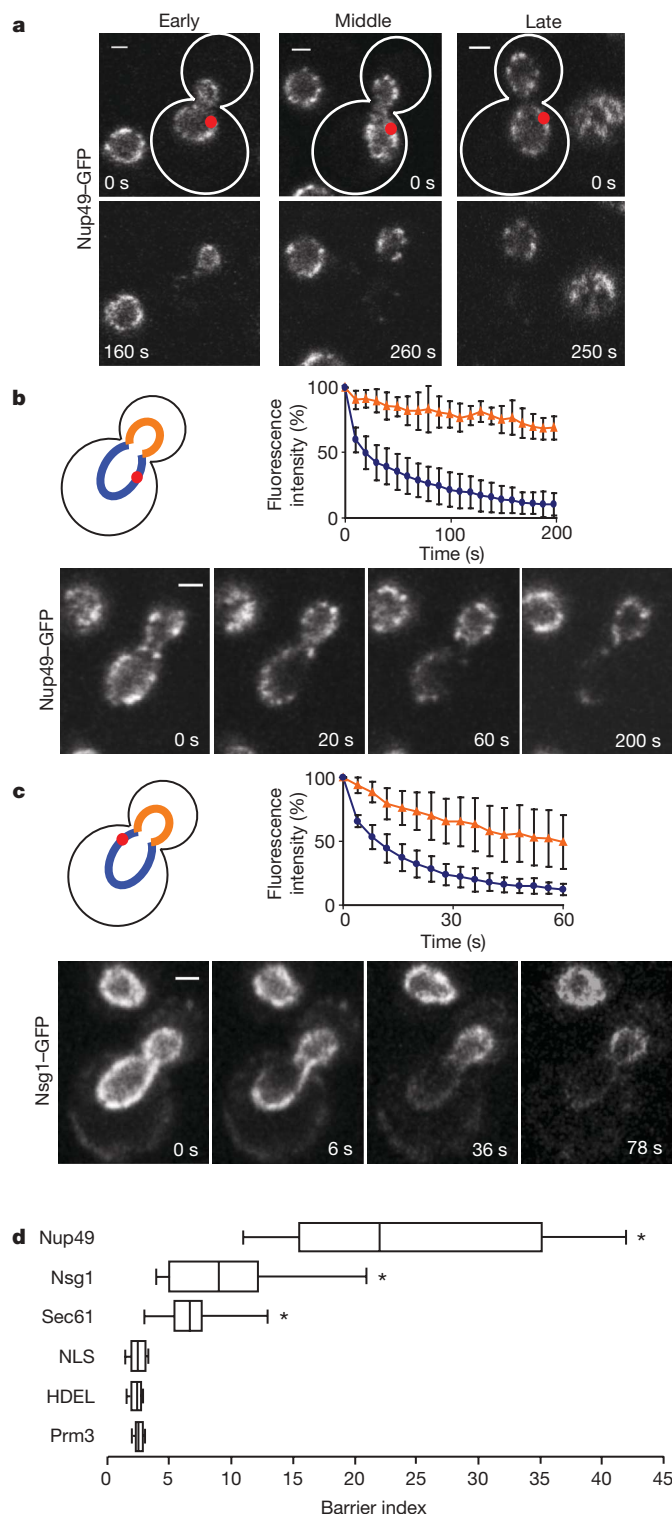


Figure 1 | Restriction of lateral diffusion through the bud neck in the outer membrane of yeast anaphase nuclei. **a**, Cells expressing Nup49-GFP were subjected to constant photobleaching (red, bleached areas) at different stages of anaphase. **b–c**, Photobleaching analysis of cells expressing Nup49-GFP ($n = 13$) or Nsg1-GFP ($n = 12$). The bleached (red) and monitored areas (orange and blue) are depicted. Graphs show average kinetics of fluorescence decay. Scale bars, 1 μm ; error bars, 1 s.d.; elapsed time is indicated in each movie frame. **d**, Distribution of the barrier index (see Methods) for the indicated proteins ($12 \leq n \leq 15$, except GFP-HDEL ($n = 5$), NLS-GFP ($n = 5$), Prm3-GFP ($n = 7$)). Boxes, twenty-fifth to seventy-fifth percentiles; whiskers extend to most extreme values; medians are indicated. * $P < 0.001$, in comparison with GFP-HDEL.

(mother) compartment than in the unbleached (bud) compartment (Fig. 1d). Intermediate values for this barrier index (BI ; see Methods) were observed for Nsg1-GFP and Sec61-GFP (9.8 ± 5.7 , $n = 12$; 6.9 ± 2.2 , $n = 15$, respectively, $P < 0.001$, in comparison with GFP-HDEL). Barrier indices were lowest and statistically indistinguishable for GFP-HDEL, NLS-GFP and Prm3-GFP (2.3 ± 0.5 , $n = 5$; 2.4 ± 0.5 , $n = 5$; 2.6 ± 0.3 , $n = 7$, respectively, $P < 0.001$, in comparison with Nup49-GFP). Thus, a diffusion barrier restricted the diffusion of outer-membrane proteins through the bud neck during anaphase, affecting larger structures, such as NPCs, most strongly. In early anaphase, exchange remained unrestricted in the inner membrane, the ER lumen and the nucleoplasm.

Pore segregation during mitosis

These results establish that NPCs are prevented from moving from the mother to the bud during anaphase. This raises the question of where the pores in the bud come from. To investigate the role of *de novo* pore insertion, we photobleached Nup49-GFP in the mother lobe of early-anaphase nuclei and monitored recovery (Fig. 2a, Supplementary Fig. 4a, Supplementary Movie 7). In eight of eight recordings, no recovery was observed in the mother for at least 10 min. Thus, no pores were inserted *de novo* in the mother. During this time, however, fluorescence increased in five of the eight corresponding buds (Fig. 2a, average trace in Supplementary Fig. 4a). Because no fluorescence could come from the mother nucleus, fluorescence increase in the bud was probably due to *de novo* insertion. Accordingly, when we photobleached the bud fraction of similar nuclei, substantial recovery (average recovery, $69\% \pm 29\%$) was observed in all buds ($n = 14$) within less than 5 min (Fig. 2b, average trace in Supplementary Fig. 4b). Little to no fluorescence was lost in the corresponding mothers (average loss, $14\% \pm 11\%$; Supplementary Fig. 4c). Thus, pore insertion might account for the fluorescence in the bud. Accordingly, Nup49-GFP fluorescence was 2.3 ± 0.9 ($n = 31$) times more intense in mother nuclei of telophase cells than in bud nuclei (Fig. 2c, d) and the pore assembly mutations *nic96-1* (ref. 23) and *nup133Δ* (ref. 24) strongly increased this bias (in *nic96-1* at restrictive temperature 30°C , mother:bud ratio of 5.6 ± 4.2 , $n = 23$, $P < 0.001$, in comparison with wild type; in *nup133Δ* at 22°C , 7.3 ± 6.4 , $n = 25$, $P < 0.001$, in comparison with wild type; in wild type at 30°C , 2.2 ± 0.5 , $n = 26$; in wild type at 22°C , 2.3 ± 0.9 , $n = 31$; Fig. 2c, d). Thus, NPCs in the mother were mainly the pre-existing pores, whereas those in the bud nucleus largely depended on *de novo* insertion. Furthermore, quantification of fluorescence distribution in the *nic96-1* cells indicated that on average only 15% of pre-existing pores passed into the bud. This establishes that a diffusion barrier retains pre-existing outer-membrane molecules in the mother during anaphase, whereas *de novo* insertion of NPCs is biased towards the bud. Therefore, the division of the yeast nucleus is asymmetric with respect to the age of the pores, pre-existing material being preferentially segregated to the mother.

The diffusion barrier on the nuclear envelope is septin dependent

To address the physiological relevance of this process, we next sought to perturb the barrier. Because of its position, we first investigated whether bud neck proteins contributed to barrier establishment or maintenance. All known bud neck structures depend on septins, which assemble into a filamentous, membrane-associated collar at the bud neck²⁵. Using the temperature-sensitive septin allele *cdc12-6* (ref. 26), we found that septin function is required for nuclear compartmentalization (in *cdc12-6* at 22°C , $BI_{\text{Nsg1-GFP}} = 7.9 \pm 2.9$, $n = 9$; in wild type at 22°C , 9.8 ± 5.7 , $n = 12$, $P > 0.05$; in *cdc12-6* at 37°C , 4.4 ± 1.7 , $n = 11$, $P < 0.01$; Fig. 3a, b, d, Supplementary Movie 8). Significant compartmentalization defects were also found in the cells lacking the bud neck protein Bud6 (ref. 27; Supplementary Movie 9) and the non-essential septin Shs1 (ref. 28), whereas the Bud6-associated formin Bnr1 (ref. 29) contributed less ($BI_{\text{Nsg1-GFP}} = 5.4 \pm 1.5$, $n = 11$, $P < 0.05$; 5.9 ± 3.1 , $n = 13$, $P < 0.05$; 7.9 ± 2.9 , $n = 15$,

$P > 0.05$, respectively; Fig. 3c, d, Supplementary Fig. 4d, e). Accordingly, the nucleoporin Nup49–GFP exchanged more rapidly between mother and bud in the early-anaphase *bud6Δ* and *cdc12-6* mutant cells, in comparison with wild type ($BI_{\text{Nup49-GFP}} = 16.0 \pm 4.5$, $n = 12$; 15.0 ± 7.1 , $n = 13$; 24.8 ± 10.7 , $n = 13$, respectively, $P < 0.05$; Supplementary Fig. 5a). Moreover, the telophase distribution of NPCs between mother and bud was close to symmetric in the *bud6Δ* mutant cells (in *bud6Δ*, mother:bud

ratio of Nup49–GFP fluorescence of 1.4 ± 0.5 , $n = 25$, in wild type, 2.3 ± 0.9 , $n = 31$, $P < 0.001$; Fig. 2c, d, Supplementary Fig. 5b, c). Thus, formation or maintenance of the diffusion barrier in the nuclear envelope depended on Bud6 and the septin collar.

Model non-centromeric plasmid is anchored to nuclear pores

We next investigated whether the nuclear diffusion barrier contributes to the asymmetric division of yeast nuclei. The plasmid pPCM14 (ref. 19), which carries 224 *tetO* repeats and an excisable centromere, was introduced into cells expressing TetR–GFP constitutively and a recombinase under the control of a galactose-inducible promoter. In glucose-containing media, the centromere ensured stability of the plasmid, which segregated symmetrically (Fig. 4a). After 4 h in galactose, over 90% of plasmids excised their centromere (also reported in ref. 19), moved randomly and remained in the mother nucleus at anaphase (Figs 4a, 5a, b). Visualization of the nuclear envelope using Nup49–CFP indicated that non-centromeric (*CEN*[−]) pPCM14 was localized with high frequency at the nuclear periphery (equatorial sections through nuclei, Fig. 4b). In 71% of the cases ($n > 200$), the distance between the plasmid and the nuclear rim could not be resolved and was therefore below $0.2 \mu\text{m}$ (Fig. 4c). Only 10% of the plasmids were more than $0.3 \mu\text{m}$ away from the rim, consistent with the distribution of ERCs and other plasmids containing

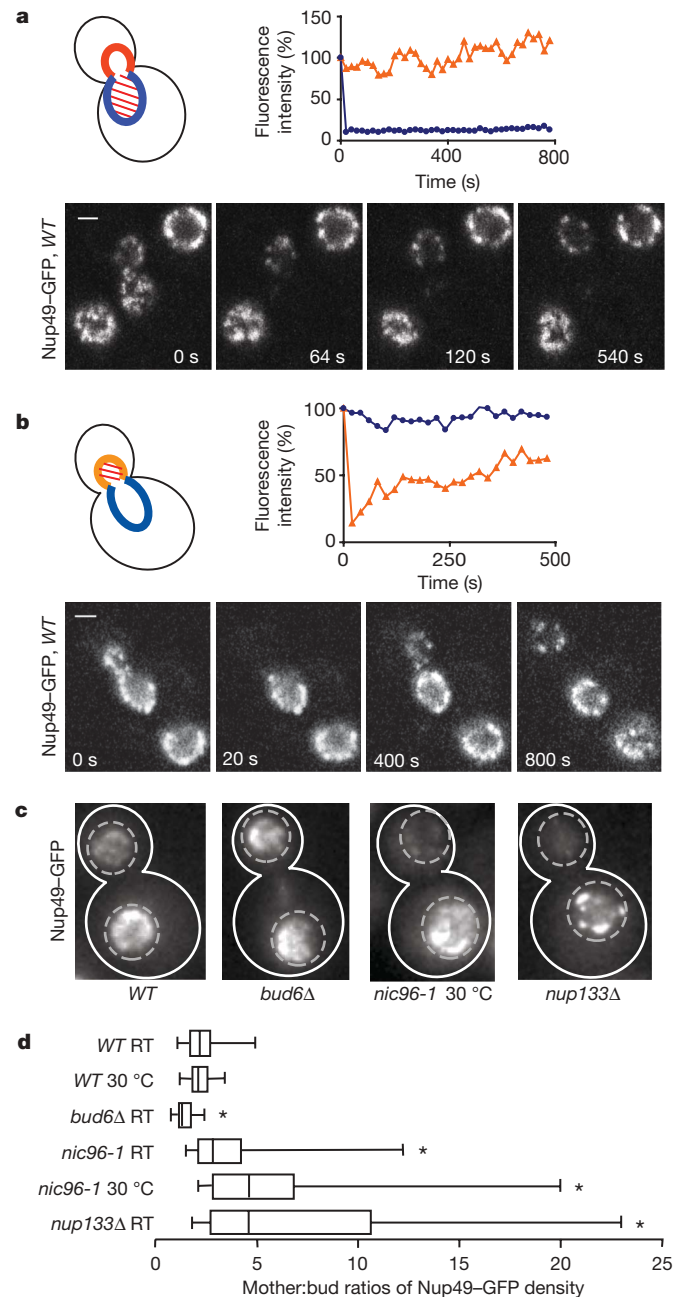


Figure 2 | A diffusion barrier in the nuclear envelope retains pre-existing nuclear pores in the mother cell. **a, b**, Fluorescence recovery after photobleaching (FRAP) in Nup49–GFP. The bleached regions are hatched in red. Each graph shows kinetics of fluorescence recovery in the two measured regions in the corresponding movie frames. Scale bars, $1 \mu\text{m}$; elapsed time is indicated in each movie frame. **c**, Fluorescence micrographs of telophase cells of the indicated genotype, expressing Nup49–GFP. Solid outlined regions are contours of the cells; dashed outlined regions are contours of the nuclei. **d**, Mean mother nucleus:bud nucleus Nup49–GFP intensity for telophase cells of the indicated phenotype. Boxes and whiskers as in Fig. 1d. For each strain, $23 \leq n \leq 31$. * $P < 0.001$, in comparison with wild type (WT) at the corresponding temperature. RT, room temperature (22°C).

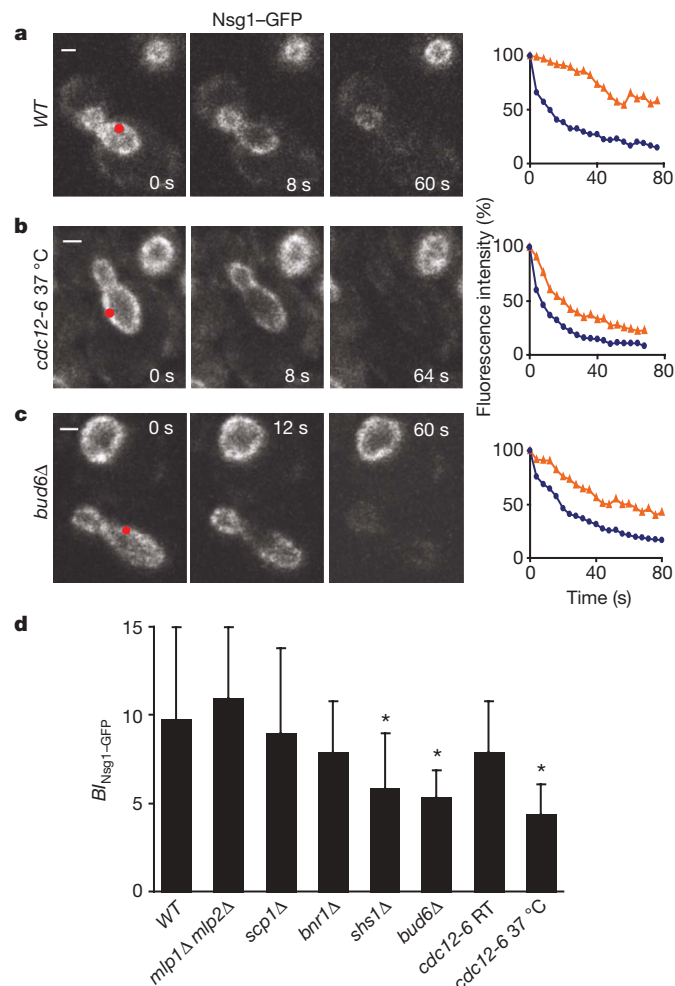
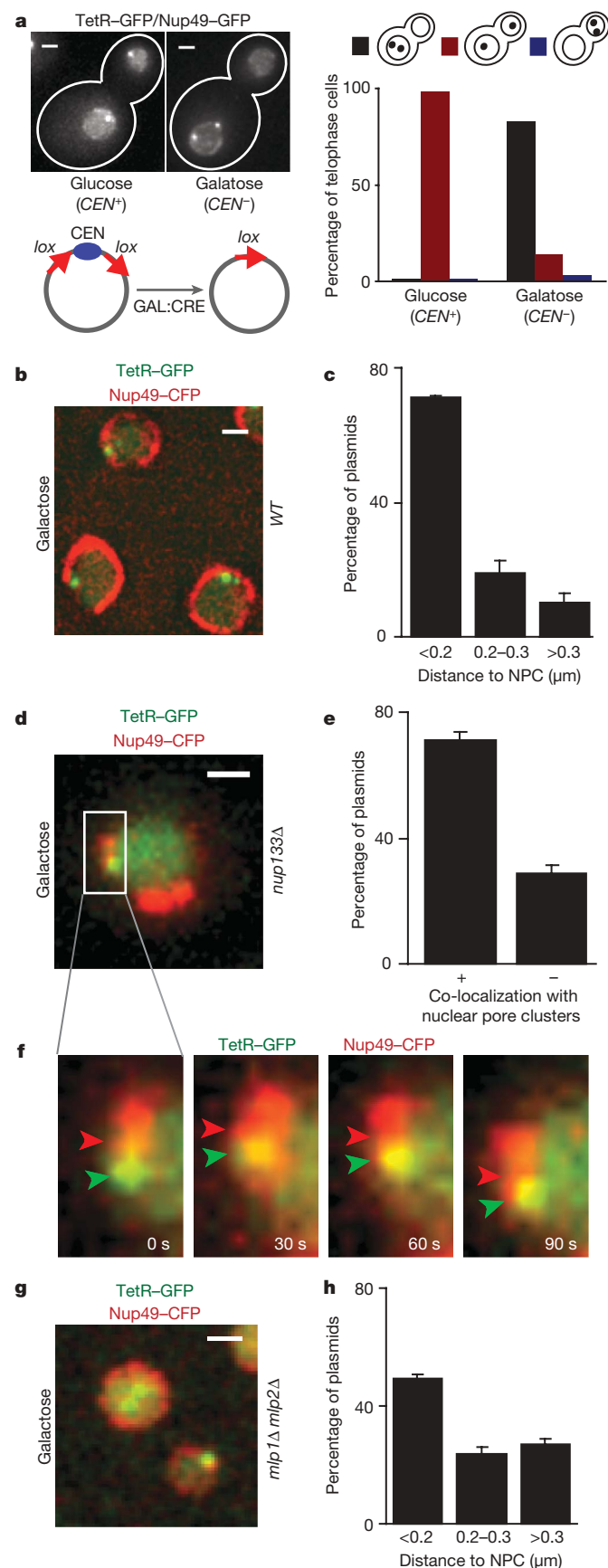


Figure 3 | Nuclear diffusion barrier is septin dependent and Bud6 dependent. **a–c**, Photobleaching analysis of cells expressing Nsg1–GFP and carrying the indicated mutations (as in Fig. 1b). Fluorescence levels in the corresponding movie frames are shown as a function of time (as in Fig. 1b). Scale bars, $1 \mu\text{m}$; elapsed time is indicated in each movie frame. **d**, $BI_{\text{Nsg1-GFP}}$ in cells carrying the indicated mutation. For each strain, $n \geq 11$. * $P < 0.05$, in comparison with wild type. Error bars, 1 s.d.

autonomously replicating sequences^{30,31}. In *nup133Δ* mutant cells, where nuclear pores cluster³², 71% of the plasmids ($n > 200$) co-localized with NPC clusters (Fig. 4d, e) and, in time-lapse movies,



the plasmids co-diffused with the pores (plasmids remained attached to pores for 75% of the time; Fig. 4f, Supplementary Movie 10). Moreover, in cells lacking the basket proteins Mlp1 and Mlp2 (ref. 33), association of the plasmid with the nuclear periphery and, hence, with pores was reduced (49% had distance to rim $< 2 \mu m$, 26% had distance to rim $> 0.3 \mu m$, $n > 200$, $P < 0.01$, in comparison with wild type; Fig. 4g, h). These data indicate that CEN^- pPCM14 segregates asymmetrically, similarly to ERCs and other non-centromeric episomes, and is anchored to pores at the nuclear periphery. Furthermore, the nuclear basket mediated the recruitment of CEN^- pPCM14 to pores.

Plasmid retention depends on the diffusion barrier and the nuclear basket

Analysis of the segregation of CEN^- pPCM14 in wild-type cells (Fig. 5a, b) and cells that divide their nuclei in the mother (Supplementary Fig. 6) indicated that the bias towards the mother was due to neither geometry nor any intrinsic asymmetry of the dividing nucleus. Instead the bud neck appeared to function as a boundary (Supplementary Information, Supplementary Fig. 6). Therefore, using time-lapse movies (Fig. 5b, c, Supplementary Movies 12, 13), we measured the frequency of plasmid retention in the mother cells in various bud neck mutants. This frequency was highest in wild-type cells (0.91 ± 0.02 , $n = 189$), was somewhat reduced in *shs1Δ* (0.85 ± 0.02 , $n = 126$, $P < 0.05$, in comparison with wild type) and *bmr1Δ* (0.87 ± 0.01 , $n = 167$, $P < 0.05$), and was lowest in *bud6Δ* (0.76 ± 0.03 , $n = 117$, $P < 0.01$) and *cdc12-6* single mutant cells (0.72 ± 0.08 at $37^\circ C$, $n = 153$, $P < 0.01$; theoretical lowest value, 0.58; Supplementary Information). Correlation analysis (Pearson's $r = 0.93$, $P < 0.01$) indicated that the efficiency of plasmid retention in the mother cell is linked to the strength of the diffusion barrier (Fig. 5d).

Consequently, we hypothesized that plasmid retention by the diffusion barrier might require their anchorage to NPCs. Consistent with this, the segregation of CEN^- pPCM14 was randomized in the *mlp1Δ mlp2Δ* double mutant (frequency of plasmid retention was 0.72 ± 0.3 , $n = 161$, $P < 0.01$, in comparison with wild type; Fig. 5d), despite the diffusion barrier in the nuclear envelope being intact ($BI_{NSG1-CFP} = 11.0 \pm 4.3$, $n = 13$, $P > 0.05$, in comparison with wild type; Figs 3d, 5d). Thus, the barrier restricted the diffusion of CEN^- pPCM14 into the bud, by means of plasmid anchorage to NPCs.

Our data show that the asymmetric division of yeast nuclei is enforced by cortical factors located at the bud neck. These factors establish a lateral diffusion barrier in the outer membrane of the nuclear envelope. This barrier then confines the bulk of the pre-existing pores and associated non-centromeric episomes into the mother nucleus. This mechanism probably applies also to ERCs, which also localize at the nuclear periphery^{31,34}, and depend on Bud6 for their retention in mother cells ($n = 63$ for *bud6Δ*, $n = 84$ for wild type, $P < 0.01$; Fig. 5e).

Figure 4 | Subnuclear distribution of the non-centromeric plasmid pPCM14. **a**, Segregation of centromeric (CEN^+) and non-centromeric (CEN^-) pPCM14 (bright dots) in telophase cells. Nup49-GFP marks the nuclear periphery. $n(CEN^+) = 168$, $n(CEN^-) = 83$. **b**, Localization of CEN^- pPCM14 (TetR-GFP, green) and nuclear pores (Nup49-CFP, red) in equatorial sections through nuclei. **c**, Distribution of plasmid-to-NPC distances (wild type). **d**, **e**, As in **b** and **c** in *nup133Δ* cells. **f**, The outlined box in **d** subjected to time-lapse microscopy. Plasmids (green arrowheads) and NPC clusters (red arrowheads) are indicated. Elapsed time is indicated in each movie frame. **g**, Localization of CEN^- pPCM14 (green) and nuclear pores (red) in *mlp1Δ mlp2Δ* mutant cells. **h**, Distribution of plasmid-to-NPC distances in *mlp1Δ mlp2Δ* nuclei. In **c**, **e** and **h**, $n > 200$. Scale bars, $1 \mu m$; error bars, 1 s.d.

Role of the barrier in bud rejuvenation

Because the *bud6Δ* mutation disrupted the asymmetry of nuclear division, we investigated whether it interfered with the segregation of age between mother and bud. We monitored the longevity of wild-type and *bud6Δ* mutant cells produced by mothers of increasing ages. A mixture of early buds (essentially first, second and third buds) was

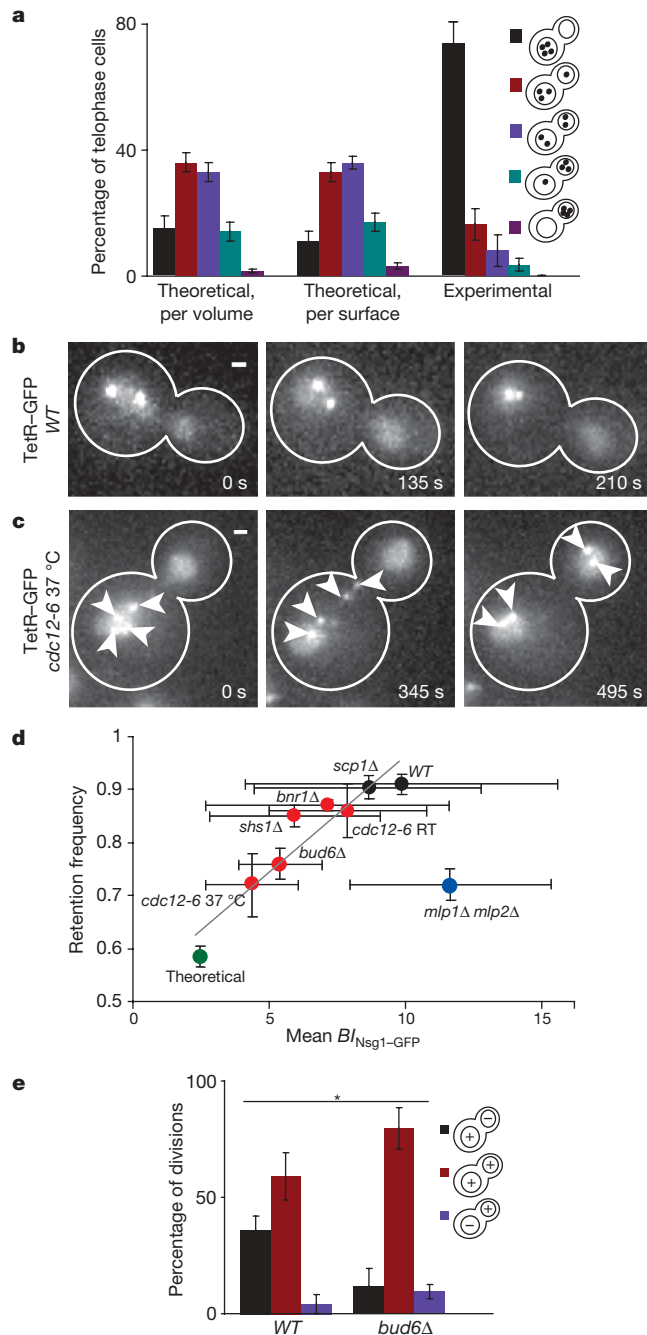


Figure 5 | Bud neck defects and plasmid segregation. **a**, Theoretical and experimental segregation of non-centromeric plasmids in wild-type cells ($n = 65$; the case of four plasmids is shown; see Supplementary Figure 6a, b for one- and two-plasmid cases). **b–c**, Movement of CEN⁺ pPCM14 (arrowheads) in cells of indicated genotype. Elapsed time is indicated in each movie frame. **d**, Correlation between median *BI_{Nsg1-GFP}* ($n \geq 11$) and plasmid retention frequency ($n \geq 117$) in bud neck (red) and NPC mutants (blue). Black, others; green, theoretical random value. **e**, Mother–bud distribution of ERC in wild-type ($n = 84$) and *bud6Δ* mutant ($n = 63$) cells, assayed as in ref. 11 (plasmid number uncontrolled). ‘+’ and ‘–’ indicate presence and absence of plasmids, respectively. * $P < 0.01$. Scale bars, 1 μm ; error bars, 1 s.d.

collected from young mothers and the number of buds that they produced before dying was monitored by pedigree analysis ($n = 130$ for *bud6Δ*, $n = 88$ for wild type). In parallel, their eighth and twelfth buds were collected and their longevity was also monitored.

In wild type, the survival curves of the early, eighth and twelfth buds were almost superimposable (median lifespan of 30 ± 0.5 generations, $P > 0.05$; Fig. 6a), indicating that they were all born similarly young. In contrast, the mortality curves of the different *bud6Δ* cells were not superimposable ($P < 0.01$; Fig. 6b). The early *bud6Δ* buds were longer lived than wild-type cells (median lifespan of 37.5 ± 1.5 generations, $P < 0.01$, in comparison with wild type), consistent with *bud6Δ* mother cells passing their ageing factors on to their progeny. In agreement with this hypothesis, *in silico* modelling of ERC-dependent ageing³⁵ suggests that loss of plasmid retention does delay ageing in early buds (Supplementary Fig. 7a). Furthermore, the buds of eighth rank showed a median lifespan reduced by seven generations relative to their mothers (the early buds), whereas the median lifespan of the twelfth buds was nine generations shorter. Thus, the older the mother they came from, the shorter a lifespan the buds had.

To determine whether this phenotype was due to a lack of rejuvenation, we analysed the lifespans of every wild-type and *bud6Δ* mother–bud pair and calculated their rejuvenation indices, defined as the number of divisions by which the daughter survived its mother, divided by its rank. Full rejuvenation should cause eighth daughters to survive their mothers by eight generations, and twelfth daughters by 12 generations. Therefore, the index should reach one upon full rejuvenation and drop to zero if there is no rejuvenation (the daughter inherits its mother’s age, and they die simultaneously). In wild type, rejuvenation index reached 0.6 (s.e.m. of 0.2, $n = 31$) for the eighth buds and 0.9 (s.e.m. of 0.2, $n = 41$) for the twelfth buds (Fig. 6c). In contrast, the rejuvenation index of the *bud6Δ* cells was to 0.1 (s.e.m. of 0.2, $n = 69$) for the eighth buds and 0.2 (s.e.m. of 0.2, $n = 48$) for the twelfth buds (average rejuvenation index of 0.14 for

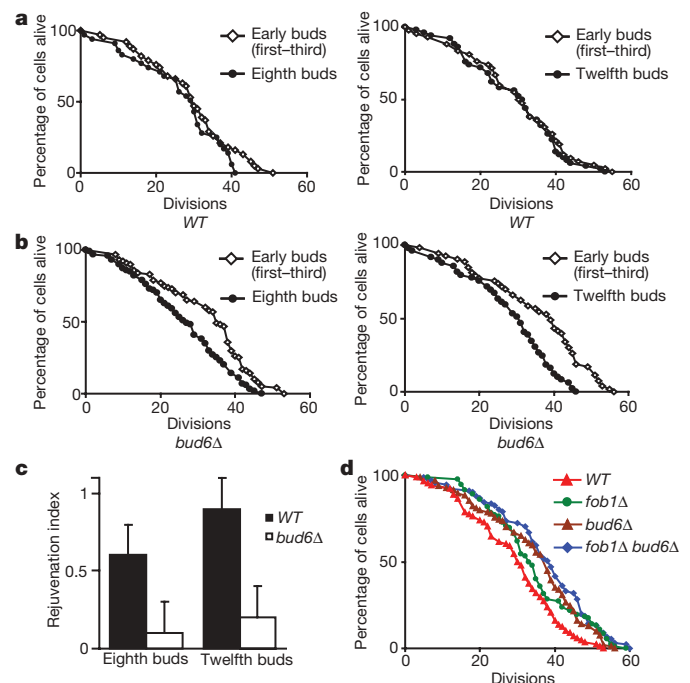


Figure 6 | Effect of diffusion barrier defects on age asymmetry. **a**, Survival curves of early (first–third) buds ($n = 38$ and 50) and their eighth ($n = 35$) and twelfth ($n = 42$) daughters, in wild type. **b**, As in **a** but in *bud6Δ* mutant cells ($n_{\text{early buds}} = 77$ and 53, $n_{\text{eighth buds}} = 71$, $n_{\text{twelfth buds}} = 49$). **c**, Rejuvenation indices for wild-type and *bud6Δ* mutant cells. Error bars, 1 s.e.m. **d**, Ageing curves of *fob1Δ* ($n = 92$), *bud6Δ* ($n = 92$) and *fob1Δ bud6Δ* ($n = 91$) mutant and wild-type ($n = 88$) early buds.

bud6Δ and 0.77 for wild type, $P < 0.01$). Thus, unlike wild type, *bud6Δ* buds failed to rejuvenate, consistent with *bud6Δ* mothers sharing their ages with their buds.

bud6Δ mutant cells no longer age through ERC accumulation

Next we investigated whether ERCs still contributed to ageing of the *bud6Δ* mutant cells. Disruption of the *FOB1* gene strongly reduces ERC formation and prolongs the lifespans of otherwise wild-type cells³⁶. Lifespan analysis of the *bud6Δ*, *fob1Δ* and *bud6Δ fob1Δ* cells indicated that in our strain background the *fob1Δ* mutant (median lifespan of 34 generations, $n = 92$) lived longer than wild type (30 generations) but for less time than the *bud6Δ* mutant cells (38 generations). Notably, *fob1Δ* mutation did not prolong the lifespans of *bud6Δ* mutant cells (median lifespan in *bud6Δ fob1Δ* of 38 generations, $n = 91$; Fig. 6d). Thus, ERCs no longer accumulated in cells lacking a functional diffusion barrier and therefore did not limit their lifespans. Conversely, barrier disruption did prolong the lifespans of cells with reduced ERC formation, suggesting that the barrier contributed to the accumulation of ageing markers other than ERCs.

Loss of the diffusion barrier in the nuclear envelope is probably not the only mechanism that increases lifespan, because the long-lived mutant *scp1Δ* (ref. 37) showed no barrier defect ($BI_{NSG1-GFP} = 9.0 \pm 4.5$, $n = 11$, plasmid retention of 0.90 ± 0.02 , $n = 142$, $P > 0.05$, in comparison with wild type; Fig. 3d). However, we conclude that the diffusion barrier in the nuclear envelope controls age segregation between mothers and buds.

Our data indicate that in early anaphase a septin- and Bud6-dependent lateral diffusion barrier compartmentalizes the nuclear envelope (see model, Supplementary Fig. 7b). Compartmentalization governed the retention of pre-existing nuclear pores in the mother and the asymmetric segregation of ageing factors associated with them. DNA circles might be only one class of these ageing factors. Thus, our data underline the importance of the nuclear envelope in ageing and establish that the dynamics of the envelope contribute substantially to the control of age segregation during yeast mitosis. Although the situation is complicated in animal cells by the full or partial dispersal of the nuclear envelope during mitosis, a link between nuclear envelope dynamics and ageing might remain, as suggested by the contribution of nuclear lamin mutations to progeria^{38,39}. Finally, the septin collar and Bud6 not only influence nuclear dynamics, but also compartmentalize the yeast cortical ER (ref. 40). Therefore, it is possible that Bud6-dependent diffusion barriers in internal membranes have a general role in confining both nuclear and cytoplasmic ageing factors to yeast mother cells.

METHODS SUMMARY

Microbiology. Yeast strains were isogenic to S288C and were grown at 22 °C unless indicated otherwise. Genetic methods⁴¹ were standard. *SEC61-GFP*, *GFP-HDEL*, *NSG1-GFP*, *NUP49-GFP*, *PRM3-GFP*, *pPCM14*, *cdc12-6* and *nic96-1* constructs and mutations were as described in refs 17, 19, 23, 26, 40, 42, 43. *NUP49-CFP* was constructed according to ref. 44.

Microscopy techniques. All microscopy techniques were as described in refs 40, 45–47. Photobleached cells were chosen on the basis of nuclear morphology. Only cells that continued anaphase throughout the experiment were quantified. The barrier index is the ratio of $t_{30\%}$ in the non-bleached compartment to $t_{30\%}$ in the bleached compartment, where $t_{30\%}$ is the earliest time by which 30% of fluorescence has been lost.

Plasmid retention. Cells grown on glucose (mid-log phase) were galactose induced for 6 h before viewing. *cdc12-6* cells were shifted to 37 °C 2 h before viewing. Average retention frequency is $(a + \sqrt{b + 4c})/3$, where a , b and c are the percentages of cells retaining all plasmids in the mother in populations of cells containing exactly one, two and, respectively, four plasmids.

Lifespan analysis. Lifespan analyses were performed as described in ref. 9. Eighth or twelfth buds were isolated from every mother alive and analysed similarly, on the same plate.

Statistics. Statistical methods used in this study are unpaired two-tailed t -tests, one-way analysis of variance tests followed by Dunnett's post tests, log rank (Mantel–Haenszel) tests, chi-squared tests and two-tailed Pearson correlation analysis.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 13 March; accepted 25 June 2008.

Published online 27 July 2008.

- Stewart, E. J., Madden, R., Paul, G. & Taddei, F. Aging and death in an organism that reproduces by morphologically symmetric division. *PLoS Biol.* **3**, doi:10.1371/journal.pbio.0030045 (2005).
- Mortimer, R. K. & Johnston, J. R. Life span of individual yeast cells. *Nature* **183**, 1751–1752 (1959).
- Pruyne, D. & Bretscher, A. Polarization of cell growth in yeast. I. Establishment and maintenance of polarity states. *J. Cell Sci.* **113**, 365–375 (2000).
- Barral, Y., Mermall, V., Mooseker, M. S. & Snyder, M. Compartmentalization of the cell cortex by septins is required for maintenance of cell polarity in yeast. *Mol. Cell* **5**, 841–851 (2000).
- Barton, A. A. Some aspects of cell division in *Saccharomyces cerevisiae*. *J. Gen. Microbiol.* **4**, 84–86 (1950).
- Breitenbach, M. et al. in *Model Systems in Aging* (eds Nyström T. & Osiewicz, H. D.) 61–96 (Springer, 2003).
- Sinclair, D., Mills, K. & Guarente, L. Aging in *Saccharomyces cerevisiae*. *Annu. Rev. Microbiol.* **52**, 533–560 (1998).
- Hartwell, L. H., Culotti, J., Pringle, J. R. & Reid, B. J. Genetic control of the cell division cycle in yeast. *Science* **183**, 46–51 (1974).
- Kennedy, B. K., Austriaco, N. R. Jr & Guarente, L. Daughter cells of *Saccharomyces cerevisiae* from old mothers display a reduced life span. *J. Cell Biol.* **127**, 1985–1993 (1994).
- Sinclair, D. A. & Guarente, L. Extrachromosomal rDNA circles – a cause of aging in yeast. *Cell* **91**, 1033–1042 (1997).
- Murray, A. W. & Szostak, J. W. Pedigree analysis of plasmid segregation in yeast. *Cell* **34**, 961–970 (1983).
- Falcon, A. A. & Aris, J. P. Plasmid accumulation reduces life span in *Saccharomyces cerevisiae*. *J. Biol. Chem.* **278**, 41607–41617 (2003).
- Lippincott-Schwartz, J., Snapp, E. & Kenworthy, A. Studying protein dynamics in living cells. *Nature Rev. Mol. Cell Biol.* **2**, 444–456 (2001).
- Deshaies, R. J. & Schekman, R. A yeast mutant defective at an early stage in import of secretory protein precursors into the endoplasmic reticulum. *J. Cell Biol.* **105**, 633–645 (1987).
- Flury, I. et al. INSIG: a broadly conserved transmembrane chaperone for sterol-sensing domain proteins. *EMBO J.* **24**, 3917–3926 (2005).
- Wente, S. R., Rout, M. P. & Blobel, G. A new family of yeast nuclear pore complex proteins. *J. Cell Biol.* **119**, 705–723 (1992).
- Beilharz, T., Egan, B., Silver, P. A., Hofmann, K. & Lithgow, T. Bipartite signals mediate subcellular targeting of tail-anchored membrane proteins in *Saccharomyces cerevisiae*. *J. Biol. Chem.* **278**, 8219–8223 (2003).
- Pelham, H. R. Recycling of proteins between the endoplasmic reticulum and Golgi complex. *Curr. Opin. Cell Biol.* **3**, 585–591 (1991).
- Megee, P. C. & Koshland, D. A functional assay for centromere-associated sister chromatid cohesion. *Science* **285**, 254–257 (1999).
- Daigle, N. et al. Nuclear pore complexes form immobile networks and have a very low turnover in live mammalian cells. *J. Cell Biol.* **154**, 71–84 (2001).
- Belgareh, N. & Doye, V. Dynamics of nuclear pore distribution in nucleoporin mutant yeast cells. *J. Cell Biol.* **136**, 747–759 (1997).
- Bucci, M. & Wente, S. R. In vivo dynamics of nuclear pore complexes in yeast. *J. Cell Biol.* **136**, 1185–1199 (1997).
- Zabel, U. et al. Nic96p is required for nuclear pore formation and functionally interacts with a novel nucleoporin, Nup188p. *J. Cell Biol.* **133**, 1141–1152 (1996).
- Walther, T. C. et al. The conserved Nup107–160 complex is critical for nuclear pore complex assembly. *Cell* **113**, 195–206 (2003).
- Faty, M., Fink, M. & Barral, Y. Septins: a ring to part mother and daughter. *Curr. Genet.* **41**, 123–131 (2002).
- Haarer, B. K. & Pringle, J. R. Immunofluorescence localization of the *Saccharomyces cerevisiae* CDC12 gene product to the vicinity of the 10-nm filaments in the mother-bud neck. *Mol. Cell Biol.* **7**, 3678–3687 (1987).
- Amberg, D. C., Zahner, J. E., Mulholland, J. W., Pringle, J. R. & Botstein, D. Aip3p/Bud6p, a yeast actin-interacting protein that is involved in morphogenesis and the selection of bipolar budding sites. *Mol. Biol. Cell* **8**, 729–753 (1997).
- Mino, A. et al. Shs1p: a novel member of septin that interacts with spa2p, involved in polarized growth in *Saccharomyces cerevisiae*. *Biochem. Biophys. Res. Commun.* **251**, 732–736 (1998).
- Imamura, H. et al. Bni1p and Bnr1p: downstream targets of the Rho family small G-proteins which interact with profilin and regulate actin cytoskeleton in *Saccharomyces cerevisiae*. *EMBO J.* **16**, 2745–2755 (1997).
- Scott-Drew, S., Wong, C. M. & Murray, J. A. DNA plasmid transmission in yeast is associated with specific sub-nuclear localisation during cell division. *Cell Biol. Int.* **26**, 393–405 (2002).
- Oakes, M. et al. Mutational analysis of the structure and localization of the nucleolus in the yeast *Saccharomyces cerevisiae*. *J. Cell Biol.* **143**, 23–34 (1998).
- Doye, V., Wepf, R. & Hurt, E. C. A novel nuclear pore protein Nup133p with distinct roles in poly(A)⁺ RNA transport and nuclear pore distribution. *EMBO J.* **13**, 6062–6075 (1994).

33. Strambio-de-Castillia, C., Blobel, G. & Rout, M. P. Proteins connecting the nuclear pore complex with the nuclear interior. *J. Cell Biol.* **144**, 839–855 (1999).
 34. Sinclair, D. A., Mills, K. & Guarente, L. Accelerated aging and nucleolar fragmentation in yeast *sgs1* mutants. *Science* **277**, 1313–1316 (1997).
 35. Gillespie, C. S. *et al.* A mathematical model of ageing in yeast. *J. Theor. Biol.* **229**, 189–196 (2004).
 36. Defossez, P. A. *et al.* Elimination of replication block protein Fob1 extends the life span of yeast mother cells. *Mol. Cell* **3**, 447–455 (1999).
 37. Gourlay, C. W. & Ayscough, K. R. A role for actin in aging and apoptosis. *Biochem. Soc. Trans.* **33**, 1260–1264 (2005).
 38. Navarro, C. L., Cau, P. & Levy, N. Molecular bases of progeroid syndromes. *Hum. Mol. Genet.* **15**, R151–R161 (2006).
 39. Scaffidi, P. & Misteli, T. Lamin A-dependent nuclear defects in human aging. *Science* **312**, 1059–1063 (2006).
 40. Luedeke, C. *et al.* Septin-dependent compartmentalization of the endoplasmic reticulum during yeast polarized growth. *J. Cell Biol.* **169**, 897–908 (2005).
 41. Guthrie, C. & Fink, G. R. *Guide to Yeast Genetics and Molecular Biology* (Academic Press, 1991).
 42. Huh, W. K. *et al.* Global analysis of protein localization in budding yeast. *Nature* **425**, 686–691 (2003).
 43. Winzler, E. A. *et al.* Functional characterization of the *S. cerevisiae* genome by gene deletion and parallel analysis. *Science* **285**, 901–906 (1999).
 44. Longtine, M. S. *et al.* Additional modules for versatile and economical PCR-based gene deletion and modification in *Saccharomyces cerevisiae*. *Yeast* **14**, 953–961 (1998).
 45. Kusch, J., Meyer, A., Snyder, M. P. & Barral, Y. Microtubule capture by the cleavage apparatus is required for proper spindle positioning in yeast. *Genes Dev.* **16**, 1627–1639 (2002).
 46. Dobbelaere, J., Gentry, M. S., Hallberg, R. L. & Barral, Y. Phosphorylation-dependent regulation of septin dynamics during the cell cycle. *Dev. Cell* **4**, 345–357 (2003).
 47. Dobbelaere, J. & Barral, Y. Spatial coordination of cytokinetic events by compartmentalization of the cell cortex. *Science* **305**, 393–396 (2004).
- Supplementary Information** is linked to the online version of the paper at www.nature.com/nature.
- Acknowledgements** We thank D. Barral and J. Sasse, P. Megee, D. Koschland, T. Lithgow, L. Guarente, V. Doye, E. Hurt and M. Winey for technical help and sharing reagents, D. Gerlich and C. Weirich for reading the manuscript, and the Barral, Meraldi and Gerlich labs for discussions. We acknowledge G. Csucs, J. Kusch and the Light Microscopy Center (LMC, ETH Zurich) for support with microscopy equipment and techniques. This work was supported by the ETH Zurich, and by a grant from the Swiss National Foundation to Y.B. and G.G.
- Author Contributions** Z.S. did most of the experiments and contributed to the data analysis and to the writing of the paper; S.B. contributed to the ageing experiments; S.B.F. contributed to the FLIP experiments; G.G. did the *in silico* modelling; and Y.B. had the idea for the project, contributed to the data analysis and wrote the paper.
- Author Information** Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to Y.B. (yves.barral@bc.biol.ethz.ch).

METHODS

Strains, plasmids and growth conditions. Yeast strains were constructed according to standard genetic techniques⁴¹ and are isogenic to S288C. Cultures were grown at room temperature (22 °C) unless otherwise indicated. The Sec61–GFP, GFP–HDEL, Nsg1–GFP and Nup49–GFP reporter constructs have been described in refs 18, 42 (<http://yeastgfp.ucsf.edu>). Nup49–CFP was constructed using the PCR-based integration system⁴⁴. Prm3–GFP (ref. 17) was expressed from a centromeric vector under control of *MET25* (also known as *MET17*) promoter (a gift from T. Lithgow, University of Melbourne, Parkville, Australia). Deletion strains were obtained from the EUROSCARF deletion collection⁴³ (<http://web.uni-frankfurt.de/fb15/mikro/euroscarf/>). The *nic96-1* (ref. 23) and *cdc12-6* mutation has been described in ref. 26. Visualization of the pPCM14 plasmid was carried out as described in ref. 19. The strain containing the system for plasmid visualization *in vivo* (a gift from P. Megee and D. Koschland, Carnegie Institution of Washington, Baltimore, USA) is W303 and was backcrossed four times into S288C background. The artificial ERC construct has been described in ref. 10 (a gift from D. Sinclair and L. Guarente, Massachusetts Institute of Technology, Cambridge, USA).

FLIP and FRAP experiments. For photobleaching experiments, cells grown on YPD plates were resuspended in non-fluorescence medium¹⁴, and immobilized on an agar pad containing non-fluorescent medium and 1.6% agar. The cells were imaged on a confocal, LSM510 microscope (Zeiss Microimaging, Germany) using a Plan-Apochromat ×63 objective, typically using 3% of the laser intensity of an argon laser (488 nm) at 35% output. Photobleaching was applied on the selected area as indicated in the figures. Bleaching pulses were iterated at 50% of the laser intensity at 35% output; the number of iterations varied depending on the photobleached protein fusion (60 iterations for NLS–GFP and GFP–HDEL, 80 iterations for Nsg1–GFP, 100 iterations for Nup49–GFP and Sec61–GFP). For FRAP, 30 iterations at the same conditions were applied to the photobleached area after the first scan. Subsequent scans were made every 20 s.

For the *cdc12-6* mutant, the cells were grown on YPD plates to mid-log phase at 22 °C (a permissive temperature) and subjected to FLIP analysis 20 min after a temperature shift to 37 °C (a restrictive temperature). All pictures of FLIP and FRAP experiments shown in the figures were treated to account for the photobleaching due to picture acquisition, whereas the movies were left untreated.

Quantification of FLIP and FRAP experiments and barrier index definition. Quantification of the fluorescent signals before and after photobleaching was performed using the software ImageJ 1.29 (<http://rsb.info.nih.gov/ij/>). For FLIP experiments, the loss of fluorescence over time was measured in the part of the dividing nucleus that had passed the bud neck and in the part that had not, in a neighbouring cell (to control for loss of fluorescence due to image acquisition), and in the background. The intensity in the region of interest was calculated as (region – background)/(control cell – background).

The barrier index was defined as the ratio of $t_{30\%}$ in the non-bleached compartment to $t_{30\%}$ in the bleached compartment, where $t_{30\%}$ is the earliest time by which 30% of the fluorescence intensity has been lost. Unpaired two-tailed *t*-tests or one-way analysis of variance tests followed by Dunnett's post test were used to assess the statistical significance of the FLIP data shown in Figs. 1, 3 and Supplementary Figs 3–5.

Other microscopy. TetR–GFP, indicating the localization of the plasmid pPCM14; Nup49–GFP, indicating nuclear pore density; Nup49–CFP in

nup133Δ mutant cells, as a reporter on nuclear pore clusters; and Nsg1–GFP, as a reporter on the nuclear envelope in wild-type and spindle-positioning mutants were visualized using TillVision software (Till Photonics) on an Olympus BX50 fluorescence microscope equipped with a piezo motor. For co-localization of pPCM14 and nuclear pores in wild type, images were acquired using Softwox software (Applied Precision) on a DeltaVision microscope (DeltaVision Spectris, Applied Precision) equipped with a piezo motor. In both cases, a ×100 objective of numerical aperture 1.4 was used.

To assess Nup49–GFP intensities in mother and bud nuclei of telophase cells, we measured mean fluorescence intensity in additive projections of 11 Z-stacks sampling entire nuclei. In experiments at restrictive temperature, wild-type and *nic96-1* cells were grown at 22 °C, shifted to 30 °C for 1 h and subjected to microscopy. Unpaired two-tailed *t*-tests were used to evaluate differences in mean intensities in wild-type and mutant cells. Ratios between nuclei sizes in telophase cells in wild-type and mutant cells were compared using chi-squared tests.

For plasmid-to-NPC co-localization in *nup133Δ* mutant cells, only nuclei with small nuclear pore clusters (less than 25% of the equatorial circumference) were counted.

Plasmid retention assay. Cells were grown on selective, glucose-containing medium to mid-log phase, shifted to YP medium containing 2% galactose as sole carbon source for at least 6 h, and subsequently subjected to bright-field microscopy. The *cdc12-6* mutant cells were shifted to 37 °C after 4 h on galactose at 22 °C, and subjected to microscopy 3 h after the temperature shift.

Plasmid distribution in telophase cells containing one, two or four plasmids was analysed. Average retention probability was calculated as $(a + \sqrt[4]{b + 4\sqrt{c}})/3$, where *a*, *b* and *c* are the percentages of cells retaining all plasmids in the mother in populations of cells containing exactly one, two and, respectively, four plasmids. Theoretical and experimental distributions of the plasmids between mother and bud were compared using chi-squared tests. Retention coefficients were compared by one-way analysis of variance tests followed by Dunnett's post tests.

Lifespan analysis. Lifespan analyses were performed as described in ref. 9, with the following modifications. Cells were thawed from stocks and grown on freshly made YPD plates two days before the beginning of the experiment, then re-struck to fresh YPD plates. Virgin buds were isolated and allowed to produce offspring. Eighth or twelfth buds were isolated from every mother alive, positioned on the same plate and subjected to the same analysis. The survival curves were compared using log rank (Mantel–Haenszel) tests.

In silico modelling. *In silico* modelling was carried out as in ref. 35, except the bias of plasmid segregation was changed towards the mother as indicated in Supplementary Figure 6a.

Statistics. One-way analysis of variance tests followed by Dunnett's post tests were used to compare barrier indices, retention coefficients and mother:bud ratios of signal density (Fig. 2d). Unpaired two-tailed *t*-tests were used to evaluate differences in median lifespans and rejuvenation indices. Survival curves (Fig. 6a, b, d) were compared using log rank (Mantel–Haenszel) tests. Chi-squared tests were used to compare ratios between nuclear sizes, theoretical and experimental distributions of plasmids, ERC distribution in wild-type and *bud6Δ* mutant cells and intranuclear distribution of plasmids in wild-type and *mlp1Δ mlp2Δ* mutant cells. Two-tailed Pearson correlation analysis was used to quantify the relationship between barrier indices and retention coefficients.

Clumps and streams in the local dark matter distribution

J. Diemand¹, M. Kuhlen², P. Madau¹, M. Zemp¹, B. Moore³, D. Potter³ & J. Stadel³

In cold dark matter cosmological models^{1,2}, structures form and grow through the merging of smaller units³. Numerical simulations have shown that such merging is incomplete; the inner cores of haloes survive and orbit as ‘subhaloes’ within their hosts^{4,5}. Here we report a simulation that resolves such substructure even in the very inner regions of the Galactic halo. We find hundreds of very concentrated dark matter clumps surviving near the solar circle, as well as numerous cold streams. The simulation also reveals the fractal nature of dark matter clustering: isolated haloes and subhaloes contain the same relative amount of substructure and both have cusped inner density profiles. The inner mass and phase-space densities of subhaloes match those of recently discovered faint, dark-matter-dominated dwarf satellite galaxies^{6–8}, and the overall amount of substructure can explain the anomalous flux ratios seen in strong gravitational lenses^{9,10}. Subhaloes boost γ -ray production from dark matter annihilation by factors of 4 to 15 relative to smooth galactic models. Local cosmic ray production is also enhanced, typically by a factor of 1.4 but by a factor of more than 10 in one per cent of locations lying sufficiently close to a large subhalo. (These estimates assume that the gravitational effects of baryons on dark matter substructure are small.)

The cold dark matter (CDM) model has been remarkably successful at describing the large-scale mass distribution of our Universe from the hot Big Bang to the present. However, the nature of the dark matter particle is best tested on small scales, where its interaction properties manifest themselves by modifying the structure of galaxy haloes and their substructures. CDM theory predicts that the growth of cosmic structures begins early, on Earth-like mass scales^{11,12}, and continues from the bottom up until galaxy clusters form that are 20 orders of magnitude more massive. Resolving small-scale structures is extremely challenging, as the range of lengths, masses, and timescales that need to be simulated is immense. We have performed the highest precision calculation—which we name Via Lactea II—of the assembly of the Galactic CDM halo. The simulation follows the growth of a Milky Way-size system from redshift 104.3 to the present. It provides the most accurate predictions on the small-scale clustering of dark matter so far available and puts constraints on the local subhalo abundance and properties. We used the parallel tree-code PKDGRAV2 (ref. 13) and sampled a galaxy-forming region with 1.1×10^9 particles of mass $4,100 M_\odot$ (where M_\odot denotes the mass of the Sun). Cosmological parameters were taken from Wilkinson Microwave Anisotropy Probe data¹⁴; see the Supplementary Information for more details and a comparison with our previous simulation¹⁵ of the Galactic CDM halo, Via Lactea.

The wealth of substructure that survives the hierarchical assembly process until the present epoch is clearly seen in Fig. 1: we resolve over 40,000 subhaloes within 402 kpc of the centre and find that they are distributed with approximately equal total mass in subhaloes per

decade of mass over the range $10^6 M_\odot$ – $10^9 M_\odot$. They have very high central phase-space densities ($\gtrsim 10^{-5} M_\odot \text{pc}^{-3} \text{km}^{-3} \text{s}^3$) owing to their steep inner density cusps and their relatively small internal velocity dispersions. This agrees well with the extremely high phase-space densities inferred from stellar motions within ultrafaint dwarf galaxies⁷. Our predicted inner subhalo densities ($0.4 M_\odot$ – $2.5 M_\odot \text{pc}^{-3}$ within 100 pc of centre, $7 M_\odot$ – $46 M_\odot \text{pc}^{-3}$ within 10 pc of centre) are also in excellent agreement with the observations^{6,7}. The fact that CDM theory naturally predicts a small-scale

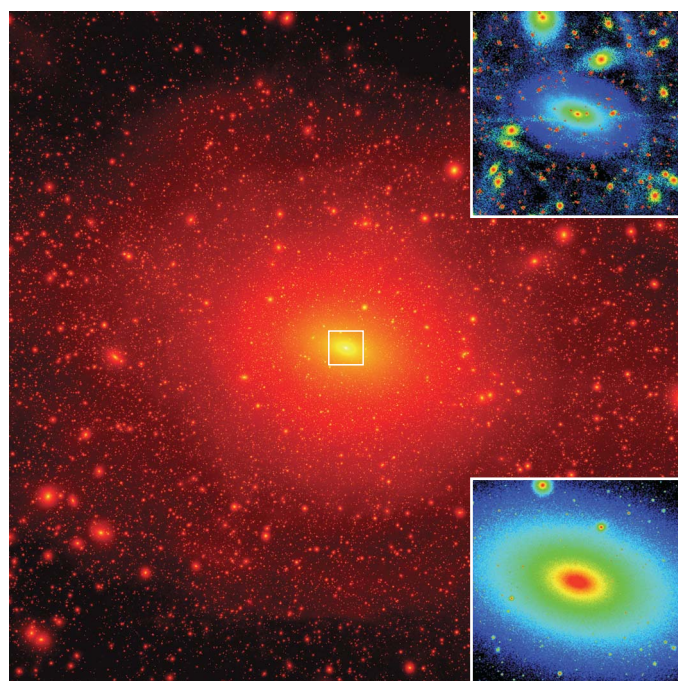


Figure 1 | Via Lactea II projected dark matter squared-density map. A cube of 800 kpc per side is shown. The insets focus on an inner cube of 40 kpc per side (outlined in white), and show local density (bottom inset) and local phase-space density calculated with EnBiD (ref. 27; top inset). The Via Lactea II simulation has a mass resolution of $4,100 M_\odot$ and a force resolution of 40 pc. It used over a million processor hours on the ‘Jaguar’ Cray XT3 supercomputer at the Oak Ridge National Laboratory. A new method was employed to assign physical, adaptive time steps¹⁹ equal to 1/16 of the local dynamical timescale (but not less than 268,000 yr), which enables the resolution of very high-density regions. Initial conditions were generated with a modified, parallel version of GRAFIC2 (ref. 28). The high-resolution region is embedded within a large periodic box (40 co-moving megaparsecs) to account for the large-scale tidal forces. The mass within $r_{200} = 402$ kpc of the centre (the radius enclosing 200 times the mean matter density) is $1.9 \times 10^{12} M_\odot$.

¹University of California, Department of Astronomy and Astrophysics, Santa Cruz, California 95064, USA. ²Institute for Advanced Study, Einstein Drive, Princeton, New Jersey 08540, USA. ³University Zurich, Institute for Theoretical Physics, Winterthurerstrasse 190, 8057 Zurich, Switzerland.

dark matter distribution that matches the observations is a real success of the model. Particle candidates that introduce a low phase-space limit, such as a sterile neutrino, or that have a high collisional cross-section, such as self-interacting dark matter, would fail these fundamental observational tests.

The phase-space map (Fig. 1, upper inset) contains coherent elongated features. These are streams which form from material removed from accreted and disrupted subhaloes. The few visible streams have quite low densities (about 100 times below the local density), but owing to their low velocity dispersion (which is about 10 times smaller than that of background particles) they can just be distinguished by their local phase-space densities (which are about $10^{-9} M_{\odot} \text{pc}^{-3} \text{km}^{-3} \text{s}^3$). These resolved streams, together with the multitude of expected finer-grained phase-space structures that we currently cannot resolve, will lead to unique signatures in direct detection experiments, especially those with directional sensitivity. In cases where the disrupted subhalo hosts a luminous satellite galaxy, the resulting streams will contain not only dark matter but also stars. This process will then produce detectable features in the Milky Way's stellar halo, like those observed in the 'field of streams'¹⁶.

Further evidence for halo substructure comes from the anomalous flux ratios in multiply imaged gravitationally lensed quasars^{17,18}. Perturbations of the light path caused by substructure can explain this phenomenon if the projected substructure fraction within 10 kpc of the centre is about 1% (refs 9,10). Within a projected distance of 10 kpc from the centre, 0.50% of the host mass belongs to resolved substructure, which could be just enough to explain the observed flux anomalies. In earlier simulations this percentage was lower; the halo simulated by Via Lactea¹⁵ predicted only 0.25%, indicating that this quantity has not yet converged in the simulations.

Via Lactea II predicts a remarkable self-similar pattern of clustering properties: our simulation integrates particle orbits extremely accurately in high-density regions¹⁹, allowing a precise determination of the density profile within the inner kiloparsecs of the Galactic halo and within the centres of its satellite galaxies. We find that a cusped profile fits the host halo's density profile well, whereas the best-fit profile with a core lies below the simulated inner densities (Fig. 2). The inner profiles of subhaloes are also consistent with the presence of cusps over their resolved ranges. They scatter around the moderate cusp index of the host halo ($\gamma = 1.24$): some of them are denser in the inner part, and some are less dense, exactly like the inner parts of field haloes, which have inner slopes of $\gamma \approx 1.2 \pm 0.2$ (ref. 20). At large radii, subhalo density profiles generally fall off faster than field halo profiles. These similarities and differences between subhalo and field halo profiles have a simple explanation: subhalo density profiles were modified by tidal mass loss, which removes material from the outside in, but does not change the inner cusp structure^{15,21}. Figure 3 shows that the dwarf satellites of the Milky Way appear to be scaled versions of the main halo not only in their inner mass distribution, but also in term of relative substructure abundances. Via Lactea II demonstrates the fractal-like appearance of the dark matter by resolving the second generation of surviving sub-substructures from the merging hierarchy. This suggest that at infinite resolution we would find a long nested series of haloes within haloes within haloes, and so forth, reminiscent of a Russian doll, all the way down the first and smallest Earth-mass haloes that form.

The multitude of dark substructures increases the dark matter annihilation signal, as it is proportional to the square of the local density. For cusped profiles (Fig. 2) with some fixed inner slope ($\gamma < 1.5$) we find the following simple scaling relation for the annihilation:

$$L \propto \rho_s^2 r_s^3 \propto V_{\text{max}}^4 / r_{V_{\text{max}}} \propto V_{\text{max}}^3 \sqrt{c_V}$$

(see the Supplementary Information for the definition of the concentration c_V and its values). In combination with the steep subhalo velocity function $N(> V_{\text{max}}) \propto V_{\text{max}}^{-3}$, this implies that subhaloes of all sizes contribute about equally to the total signal coming from the

Galactic dark halo. Taking the higher concentrations of smaller systems²² into account, we find that small subhaloes contribute more than large ones^{23,24}. Summing the values of $V_{\text{max}}^4 / r_{V_{\text{max}}}$ for all the resolved subhaloes in Via Lactea II gives a number close (97%) to the host halo's $V_{\text{max}}^4 / r_{V_{\text{max}}}$ value; that is, the resolved subhaloes already contribute as much as their smooth host would alone. In other words, the 'substructure boost factor' is at least 1.97. Extrapolating down to microsubhaloes of size $V_{\text{max}} = 0.25 \text{ m s}^{-1}$, taking into account how concentrations depend on subhalo size²² and position (see the Supplementary Information), and assuming a uniform distribution of subhalo inner slopes α between 1.0 and 1.5 leads to a total boost factor of 14.6. Most of it comes from very small clumps: halting the same extrapolation at $V_{\text{max}} = 44 \text{ m s}^{-1}$ lowers the boost factor to 6.6. Although the contribution from small, dark clumps is not affected by baryons, it may not dominate the total signal in situations in which baryonic collapse greatly increases the central dark matter densities in larger haloes. However, the net effect of stars, black holes and galaxy formation is unclear, and it may actually lead to a reduction in the central dark matter densities.

The detailed distribution of cusp indices is still unknown, because only a few haloes have been simulated with sufficient resolution²⁰. For the annihilation boost factors the existence of a few steep cusps

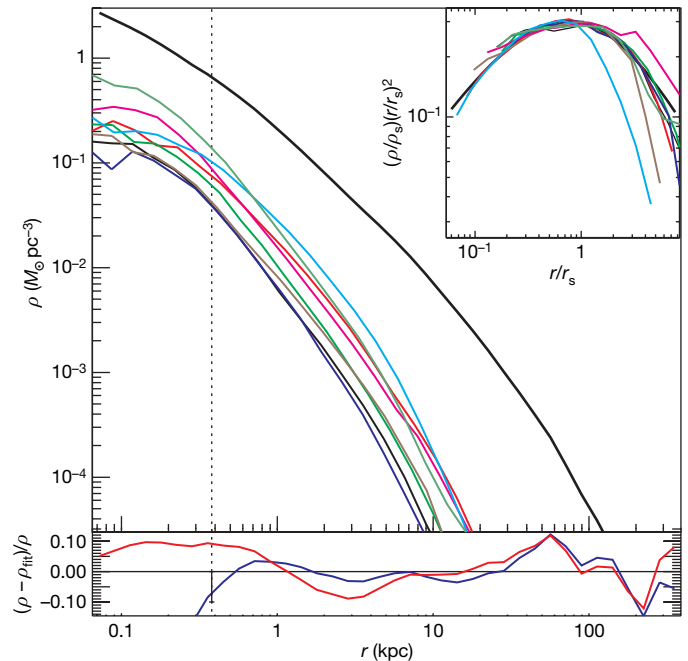


Figure 2 | Density profiles of main halo and subhaloes. Main panel, profile of the Milky Way halo (thick line) and of eight large subhaloes (thin lines). Lower panel, the relative differences between the simulated main halo profile and a fitting formula with a core²⁹, $\rho(r) = \rho_s \exp(-2/\alpha[(r/r_s)^\alpha - 1])$ (best-fit parameters $\alpha = 0.170$, $r_s = 21.5 \text{ kpc}$, $\rho_s = 1.73 \times 10^{-3} M_{\odot} \text{pc}^{-3}$; red curve), and one with a cusp²⁰, $\rho(r) = \rho_s (r/r_s)^{-\gamma} (r/r_s + 1)^{-3+\gamma}$ (best-fit inner slope $\gamma = 1.24$, $r_s = 28.1 \text{ kpc}$, $\rho_s = 3.50 \times 10^{-3} M_{\odot} \text{pc}^{-3}$; blue curve). The vertical dotted line indicates the estimated convergence radius of 380 pc: simulated local densities are only lower limits within this radius and should be correct to within 10% at this radius (vertical error bar) and outside it. The cusped profile is a good fit to the inner halo, whereas the profile with a core has too shallow a slope in the inner few kiloparsecs, causing it to overestimate densities around 4 kpc and to underestimate them at all radii less than 1 kpc. The simulated densities are higher than the best-fit profile with a core even at 80 pc, where they are certainly underestimated owing to numerical limitations. We find the same behaviour in the inner few kiloparsecs in all six snapshots that we have analysed so far between redshifts $z = 3$ and $z = 0$. The large residuals in the outer haloes, on the other hand, are transient features; they are different in every snapshot. Inset, rescaled host halo (thick line) and subhalo (thin lines) density profiles multiplied by squared radius to reduce the vertical range of the figure.

near $\gamma = 1.5$ would make a big difference, because the signal diverges logarithmically towards the centre in a $\gamma = 1.5$ cusp. Cutting the assumed uniform distribution of inner slopes at $\gamma = 1.4$ instead of $\gamma = 1.5$ gives boost factors of 9.9 instead of 14.6 and 4.3 instead of 6.6. These factors imply that most of the extra-Galactic γ -ray background from dark matter annihilation²³, which will be constrained or even detected by the upcoming Gamma-ray Large Area Space Telescope mission, should be emitted by subhaloes, and not by distinct host haloes.

As well as γ -rays, dark matter annihilation would produce charged particles and antiparticles that, owing to magnetic field entanglement, propagate over much shorter distances within the Galaxy. Space-based experiments (like the Payload for Antimatter Matter Exploration and Light-nuclei Astrophysics mission and, in the near future, the Alpha Magnetic Spectrometer experiment AMS-02) could detect antiparticles produced in dark matter annihilations within about 1 kpc of the centre²⁵. To determine what fraction of this local annihilation would happen in nearby subhaloes, we constrain the local boost factor using the same assumptions as above ($\gamma = 1$ –1.5, $V_{\max} \geq 0.25 \text{ m s}^{-1}$), but now we only include subhaloes within 1 kpc of the solar system (see the Supplementary Information for the local

subhalo abundance). The resulting signal is 40% of the smooth halo signal, giving a boost factor of 1.4, which we estimated using the spherically averaged density at 8 kpc ($\rho_0 = 0.40 \text{ GeV c}^{-2} \text{ cm}^{-3}$, where c denotes the speed of light). Explaining the positron excess from the High-Energy Antimatter Telescope experiment²⁶ with local dark matter annihilation requires boost factors of about 3 to 100 (ref. 25). When a relatively large subhalo happens to lie within 1 kpc of the Solar System, the higher boost factors can be achieved without violation of the local subhalo constraints from our simulation. Such cases are possible, but are not likely to occur: only 5.2% of all random realizations have a boost factor ≥ 3 (caused by there being a clump with $V_{\max} \geq 3.4 \text{ km s}^{-1}$ within 1 kpc of the Solar System). In only 1.0% of the cases, the boost factor is ≥ 10 , owing to a nearby, large- V_{\max} ($\geq 5.6 \text{ km s}^{-1}$) subhalo.

Received 1 February; accepted 27 May 2008.

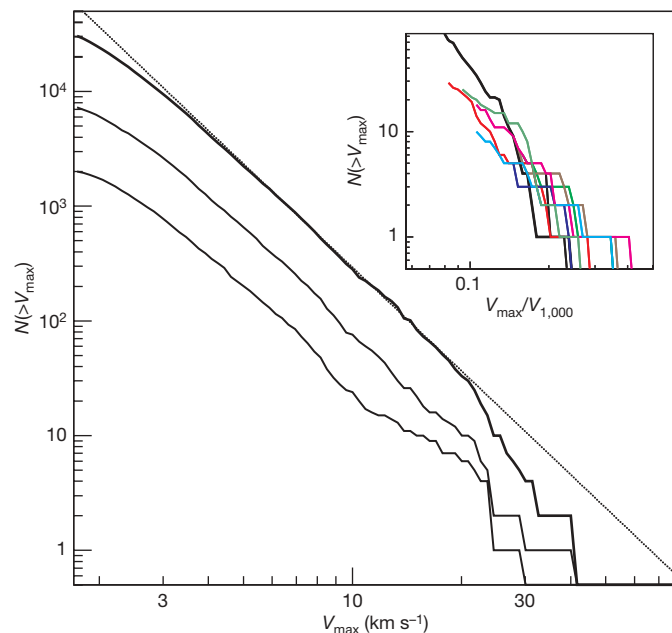


Figure 3 | Subhalo and sub-subhalo abundances. Number of subhaloes above V_{\max} within $r_{200} = 402 \text{ kpc}$ (thick solid line) and within 100 and 50 kpc of the galactic centre (thin solid lines, from top to bottom). V_{\max} is the peak height of the subhalo circular velocity $v_{\text{circ}} = \sqrt{GM(<r)/r}$, where $M(<r)$ is the mass enclosed within a sphere of radius r and G is the gravitational constant. V_{\max} serves as a simple proxy for the mass of a subhalo. The dotted line is $N(>V_{\max}) = 0.036(V_{\max}/V_{\max, \text{host}})^{-3}$, where $V_{\max, \text{host}} = 201 \text{ km s}^{-1}$ (at $r_{V_{\max, \text{host}}} = 60 \text{ kpc}$). It fits the subhalo abundance above $V_{\max} \approx 3.5 \text{ km s}^{-1}$. The number of smaller subhaloes is artificially reduced by numerical limitations. Within r_{200} this halo has 1.7 times more substructure than the first Via Lactea halo¹⁵, a factor well within the halo-to-halo scatter³⁰. Within 50 kpc this grows to a factor of 2.6, probably because of the improved mass and time resolution of Via Lactea II, which enables better resolution of inner substructure. Inset, sub-subhalo abundance within $r_{1,000}$ (enclosing 1,000 times the mean matter density) of the centres of eight large subhaloes (those in Fig. 2; thin solid lines). $r_{1,000}$ is well inside the tidal radius for these systems. The thick solid line shows the subhalo abundance of the host halo inside its $r_{1,000}$, which equals 213 kpc. The subhalo and sub-subhalo V_{\max} values are given in units of the $V_{1,000} = \sqrt{GM(<r_{1,000})/r_{1,000}}$ of the corresponding host halo or, respectively, subhalo. Lines stop at $V_{\max} = 2 \text{ km s}^{-1}$. The mean sub-substructure abundance is consistent with the scaled-down version of the main halo, and both the mean abundance and the scatter agree with results in ref. 30 for distinct field haloes.

1. Peebles, P. J. E. Large-scale background temperature and mass fluctuations due to scale-invariant primeval perturbations. *Astrophys. J.* **263**, L1–L5 (1982).
2. Blumenthal, G. R., Faber, S. M., Primack, J. R. & Rees, M. J. Formation of galaxies and large-scale structure with cold dark matter. *Nature* **311**, 517–525 (1984).
3. White, S. D. M. & Rees, M. J. Core condensation in heavy halos – A two-stage theory for galaxy formation and clustering. *Mon. Not. R. Astron. Soc.* **183**, 341–358 (1978).
4. Ghigna, S. *et al.* Dark matter haloes within clusters. *Mon. Not. R. Astron. Soc.* **300**, 146–162 (1998).
5. Klypin, A., Kravtsov, A. V., Valenzuela, O. & Prada, F. Where are the missing galactic satellites? *Astrophys. J.* **522**, 82–92 (1999).
6. Strigari, L. E. *et al.* The most dark matter dominated galaxies: Predicted gamma-ray signals from the faintest milky way dwarfs. *Astrophys. J.* **678**, 614–620 (2008).
7. Simon, J. D. & Geha, M. The kinematics of the ultra-faint milky way satellites: Solving the missing satellite problem. *Astrophys. J.* **670**, 313–331 (2007).
8. Belokurov, V. *et al.* Cats and dogs, hair and a hero: A quintet of new milky way companions. *Astrophys. J.* **654**, 897–906 (2007).
9. Dalal, N. & Kochanek, C. S. Direct detection of cold dark matter substructure. *Astrophys. J.* **572**, 25–33 (2002).
10. Metcalf, R. B., Moustakas, L. A., Bunker, A. J. & Parry, I. R. Spectroscopic gravitational lensing and limits on the dark matter substructure in Q2237+0305. *Astrophys. J.* **607**, 43–59 (2004).
11. Green, A. M., Hofmann, S. & Schwarz, D. J. The power spectrum of SUSY-CDM on subgalactic scales. *Mon. Not. R. Astron. Soc.* **353**, L23–L27 (2004).
12. Diemand, J., Moore, B. & Stadel, J. Earth-mass dark-matter haloes as the first structures in the early Universe. *Nature* **433**, 389–391 (2005).
13. Stadel, J. G. *Cosmological N-body Simulations and Their Analysis*. PhD thesis, Univ. Washington (2001).
14. Spergel, D. N. *et al.* Three-year Wilkinson Microwave Anisotropy Probe (WMAP) observations: Implications for cosmology. *Astrophys. J. Suppl. Ser.* **170**, 377–408 (2007).
15. Diemand, J., Kuhlen, M. & Madau, P. Formation and evolution of galaxy dark matter halos and their substructure. *Astrophys. J.* **667**, 859–877 (2007).
16. Belokurov, V. *et al.* The field of streams: Sagittarius and its siblings. *Astrophys. J.* **642**, L137–L140 (2006).
17. Mao, S. & Schneider, P. Evidence for substructure in lens galaxies? *Mon. Not. R. Astron. Soc.* **295**, 587–594 (1998).
18. Metcalf, R. B. & Madau, P. Compound gravitational lensing as a probe of dark matter substructure within galaxy halos. *Astrophys. J.* **563**, 9–20 (2001).
19. Zemp, M., Stadel, J., Moore, B. & Carollo, C. M. An optimum time-stepping scheme for N-body simulations. *Mon. Not. R. Astron. Soc.* **376**, 273–286 (2007).
20. Diemand, J., Moore, B. & Stadel, J. Convergence and scatter of cluster density profiles. *Mon. Not. R. Astron. Soc.* **353**, 624–632 (2004).
21. Kazantzidis, S. *et al.* Density profiles of cold dark matter substructure: Implications for the missing-satellites problem. *Astrophys. J.* **608**, 663–679 (2004).
22. Bullock, J. S. *et al.* Profiles of dark haloes: evolution, scatter and environment. *Mon. Not. R. Astron. Soc.* **321**, 559–575 (2001).
23. Ullio, P., Bergström, L., Edsjö, J. & Lacey, C. Cosmological dark matter annihilations into γ rays: A closer look. *Phys. Rev. D* **66**, 123502 (2002).
24. Colafrancesco, S., Profumo, S. & Ullio, P. Multi-frequency analysis of neutralino dark matter annihilations in the Coma cluster. *Astron. Astrophys.* **455**, 21–43 (2006).
25. Lavalle, J., Yuan, Q., Maurin, D. & Bi, X. Full calculation of clumpiness boost factors for antimatter cosmic rays in the light of Λ CDM N-body simulation results. *Astron. Astrophys.* **479**, 427–452 (2008).
26. Beatty, J. J. *et al.* New measurement of the cosmic-ray positron fraction from 5 to 15 GeV. *Phys. Rev. Lett.* **93**, 241102 (2004).
27. Sharma, S. & Steinmetz, M. Multidimensional density estimation and phase-space structure of dark matter haloes. *Mon. Not. R. Astron. Soc.* **373**, 1293–1307 (2006).

28. Bertschinger, E. Multiscale Gaussian random fields and their application to cosmological simulations. *Astrophys. J. Suppl. Ser.* **137**, 1–20 (2001).
29. Navarro, J. F. *et al.* The inner structure of Λ CDM haloes - III. Universality and asymptotic slopes. *Mon. Not. R. Astron. Soc.* **349**, 1039–1051 (2004).
30. Reed, D. *et al.* Dark matter subhaloes in numerical simulations. *Mon. Not. R. Astron. Soc.* **359**, 1537–1548 (2005).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements It is a pleasure to thank B. Messer and the Scientific Computing Group at the National Center for Computational Sciences for their help.

The Via Lactea II simulation was performed at the Oak Ridge National Laboratory through an award from the US Department of Energy's Office of Science as part of the 2007 Innovative and Novel Computational Impact on Theory and Experiment (INCITE) programme. Additional computations (initial conditions generation, code optimizations and smaller test runs) were carried out on the MareNostrum supercomputer at the BSC, on Columbia at NASA Ames and on the UCSC Astrophysics Supercomputer Pleiades. This work was supported by NASA and the Swiss National Science Foundation.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to J.D. (diemand@ucolick.org).

Determination of the fermion pair size in a resonantly interacting superfluid

Christian H. Schunck¹, Yong-il Shin¹, André Schirotzek¹ & Wolfgang Ketterle¹

Fermionic superfluidity requires the formation of particle pairs, the size of which varies from the femtometre scale in neutron stars and nuclei to the micrometre scale in conventional superconductors. Many properties of the superfluid depend on the pair size relative to the interparticle spacing. This is expressed in ‘BCS–BEC crossover’ theories^{1–3}, describing the crossover from a Bardeen–Cooper–Schrieffer (BCS)-type superfluid of loosely bound, large Cooper pairs to Bose–Einstein condensates (BECs) of tightly bound molecules. Such a crossover superfluid has been realized in ultracold atomic gases where high-temperature superfluidity has been observed^{4,5}. The microscopic properties of the fermion pairs can be probed using radio-frequency spectroscopy. However, previous work^{6–8} was difficult to interpret owing to strong final-state interactions that were not well understood. Here we realize a superfluid spin mixture in which such interactions have negligible influence and present fermion pair dissociation spectra that reveal the underlying pairing correlations. This allows us to determine that the spectroscopic pair size in the resonantly interacting gas is 20 per cent smaller than the interparticle spacing. These are the smallest pairs so far observed in fermionic superfluids, highlighting the importance of small fermion pairs for superfluidity at high critical temperatures⁹. We have also identified transitions from fermion pairs to bound molecular states and to many-body bound states in the case of strong final-state interactions.

The properties of pairs are revealed in a dissociation spectrum, where pair dissociation is monitored as a function of the applied energy E . The spectrum has a sharp onset at the pair’s binding energy E_b , where the fragments have zero kinetic energy, and then spreads out at higher energies. Because a radio-frequency photon has negligible momentum, the allowed momenta for the fragments reflect the Fourier transform $\Phi(k)$ of the pair wavefunction $\phi(r)$, which has a width on the order of $1/\xi$, where ξ is the pair size. Thus, the pair size can be estimated from the spectral linewidth E_w as $\xi^2 \approx \hbar^2/mE_w$, where m is the mass of the particles and \hbar is Planck’s constant h divided by 2π .

The conceptually simplest pairs in the BCS–BEC crossover are the weakly bound molecules in the BEC limit, which are described by a spatial wavefunction $\phi_m(r) \propto e^{-r/b}/r$ with a binding energy $E_b = \hbar^2/mb^2$ and a root-mean-square size of $b/\sqrt{2}$. When the molecules are dissociated into non-interacting free particles, the spectral response is $I_m \propto \sqrt{E - E_b}/E^2$, which shows a highly asymmetric line shape with a steep rise at the molecular binding energy E_b and a long ‘tail’ at higher energies^{5,10} (Fig. 1a).

This general behaviour of the dissociation spectrum also holds in the BCS limit, where pairing is a many-body effect^{5,11}. The radio-frequency dissociation process discussed below, in the limit of negligible final-state interactions, can be considered as breaking a Cooper pair into one quasi-particle and one free particle. The radio-frequency spectrum in the BCS limit¹² has an onset at $\Delta^2/2E_F$ and the same $E^{-3/2}$

energy dependence at high energy as in the BEC limit (Fig. 1b; here E_F is the Fermi energy and Δ is the gap). Because the radio-frequency excitation occurs throughout the whole Fermi sea, it is most natural to interpret the BCS state as $N/2$ pairs with condensation energy $\Delta^2/2E_F$, where N is the total number of fermions⁵.

A spectroscopic pair size can be defined from either the onset or the width of the radio-frequency spectrum as $\xi_{th}^2 = \hbar^2/2mE_{th}$ or, respectively, $\xi_w^2 = \gamma \hbar^2/2mE_w$. Here E_{th} is the onset/threshold energy, E_w is the full width at half maximum and $\gamma = 1.89$ is a numerical constant chosen for convenience (see Fig. 1 legend). The pair sizes ξ_{th} and ξ_w , which can be directly obtained from the radio-frequency spectrum, capture the evolution of the pair size from the BCS limit to the BEC limit (Fig. 1c).

Because the radio-frequency spectra have similar behaviour in both limiting cases of the BCS–BEC crossover, we would expect comparable spectra within the crossover regime. Surprisingly, the radio-frequency spectra obtained in previous radio-frequency experiments did not fit into this picture: the line shape did not show any pronounced asymmetry and the linewidth was narrow^{6–8} (Supplementary Fig. 3). These experiments could therefore not be simply interpreted in terms of pairing energy and pair size. We will show that this is caused by strong final-state interactions and transitions to bound states.

In both the previous experiments and our new ones, the fermion pairs consist of two atoms in different hyperfine states $|a\rangle$ and $|b\rangle$. The radio-frequency pulse transfers atoms in state $|b\rangle$ to an initially unoccupied third state $|c\rangle$. In addition to pair dissociation, also referred to as a ‘bound–free’ transition and characterized by the asymmetric line shape discussed above, radio-frequency spectroscopy can induce a second kind of transition to another bound state, that is, the transfer of a pair (a, b) to a pair (a, c) (also referred to as a ‘bound–bound’ transition). The latter spectra have narrow, symmetric line shapes.

Final-state effects arise when the dissociated atom in state $|c\rangle$ interacts with atoms in state $|a\rangle$. The interaction strength is measured by the dimensionless parameter $k_F a$. Here a is the s -wave scattering length; we use a_i and a_f to denote these lengths in the initial (a, b) and, respectively, final (a, c) interactions. As discussed in detail below, final-state interactions strongly affect the radio-frequency dissociation spectra^{13–15} when $|k_F a_f| > 1$. To overcome this problem, it is necessary to change the interactions in the final state without changing those in the initial state. Our solution is the realization of a new high-temperature superfluid in ^6Li using a different combination of hyperfine states for which radio-frequency excitation with reduced final-state interactions is possible (see Methods). As a result, we were able to resolve the bound–bound and bound–free contributions to the radio-frequency spectrum, and to determine the size of fermion pairs from the asymmetric fermion pair dissociation spectra.

We have taken advantage of the fact that any two-state mixture, (1, 2), (1, 3) or (2, 3), of the three lowest hyperfine states of ^6Li

¹Department of Physics, MIT–Harvard Center for Ultracold Atoms, and Research Laboratory of Electronics, MIT, Cambridge, Massachusetts 02139, USA.

(labelled in the order of increasing hyperfine energy as $|1\rangle$, $|2\rangle$ and $|3\rangle$) exhibits a broad s -wave Feshbach resonance^{16,17}. So far, all experiments with strongly interacting fermions in ^6Li have been carried out in the vicinity of the $(1, 2)$ Feshbach resonance located at $B_{12} \approx 834$ G (B_{ij} denotes the centre of the broad s -wave Feshbach resonance in the (i, j) mixture). Surprisingly, inelastic collisions including allowed dipolar relaxation are not enhanced by the $(1, 3)$ and $(2, 3)$ Feshbach resonances. We observe that at both the $(1, 3)$ and $(2, 3)$ Feshbach resonances, superfluids can be created as well (see Methods). This doubles the number of high-temperature superfluids available for experimental studies.

The newly created $(1, 3)$ superfluid is the best choice for radio-frequency spectroscopy experiments because the final-state scattering length a_f at the $(1, 3)$ resonance position $B_{13} \approx 691$ G is small and positive ($0 < k_F a_f < 1$ for typical values of k_F). Therefore the accessible final states are either a molecule of well-defined binding energy

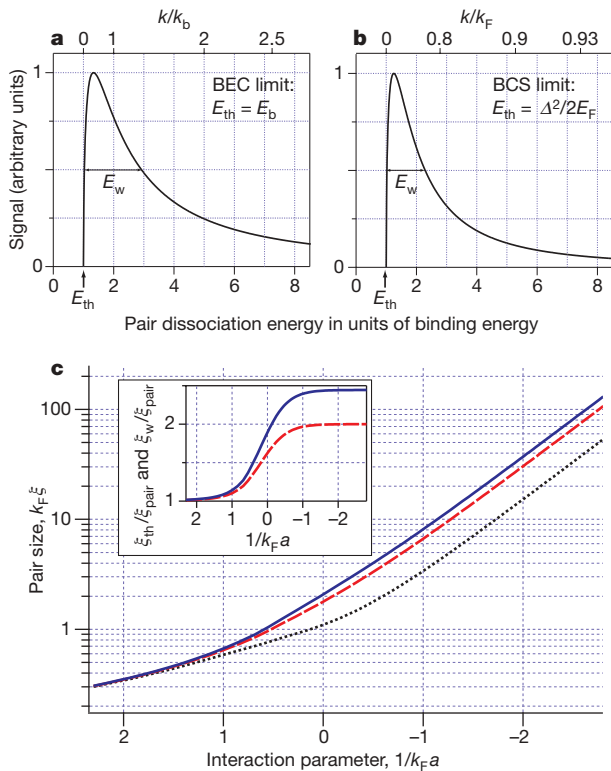


Figure 1 | Line shape of the pair dissociation spectrum in the BEC and BCS limits and the evolution of the fermion pair size in the BCS-BEC crossover^{5,11,20}. **a, b**, Simulated radio-frequency dissociation spectra in the BEC (**a**) and BCS (**b**) limits. The momentum k of the free particles after dissociation is indicated in the top axes, where $\hbar^2 k_b^2/m = E_b$ and k_F is the Fermi wavenumber. Apart from an offset, the spectra in the BEC and BCS limits have almost indistinguishable line shapes. The molecular dissociation line shape I_m with an additional offset parameter can therefore serve as a generic, model-independent fitting function for pair dissociation spectra (see Methods and Supplementary Fig. 2). **c**, The fermion pair sizes ξ_w (solid blue) and ξ_{th} (dashed red) are displayed as a function of the interaction parameter $1/k_F a$. Also shown is the two-particle correlation length ξ_{pair} (dotted black), given by $\xi_{pair} = \sqrt{\langle \phi | r^2 | \phi \rangle / \langle \phi | \phi \rangle}$, where $\phi(r) = \langle \psi | \Psi_\alpha^\dagger(r) \Psi_\beta^\dagger(0) | \psi \rangle$. Here ψ is the generalized BCS wavefunction and α and β refer to the two components². In the BEC limit, the value for the molecular size is $\xi_m = b/\sqrt{2} = \xi_{pair}$ and we choose $\gamma = 1.89$ in the definition of ξ_w , so $\xi_m = \xi_{th} = \xi_w$. In the BCS limit, $\xi_{pair} = (\pi/2\sqrt{2})\xi_c$, where $\xi_c = \hbar^2 k_F / \pi m \Delta$ is the Pippard coherence length, and we have $\xi_{th} = 2\xi_{pair}$ and $\xi_w = 2.44\xi_{pair}$. The inset shows the ratios ξ_w/ξ_{pair} (solid blue) and ξ_{th}/ξ_{pair} (dashed red). Although ξ_{pair} changes by orders of magnitude, ξ_{th} and ξ_w show the same behaviour as ξ_{pair} , deviating from each other by not more than 22%. This illustrates that the pair size can be reliably determined from the radio-frequency dissociation spectrum throughout the whole BCS-BEC crossover.

or two free, only weakly interacting atoms. The actual final-state interactions depend on whether we drive the radio-frequency transitions from $|1\rangle$ to $|2\rangle$ or from $|3\rangle$ to $|2\rangle$, allowing the comparison of spectra taken from the same sample but with different a_f values (see Methods and Supplementary Information). After preparing the $(1, 3)$ superfluid, a radio-frequency pulse resonant with the $|3\rangle$ -to- $|2\rangle$ transition is applied. Then either the losses in state $|3\rangle$ or the atoms transferred to state $|2\rangle$ are monitored (see Methods). All spectra are plotted versus frequency or energy relative to the atomic resonance, that is, relative to the energy E_0 required to transfer an atom from $|3\rangle$ to $|2\rangle$ in the absence of atoms from state $|1\rangle$.

The main results of this paper are the spectra observed in the $(1, 3)$ BEC-to-BCS crossover between 670 and 710 G (Fig. 2). The spectra have the asymmetric shape characteristic of pair dissociation and are indeed well fit by a generic pair dissociation line shape convolved with the line shape of the square excitation pulse (see Fig. 1 and Methods). If the frequency axis is scaled by E_w and the spectra are shifted to show the same onset, all three spectra overlap as shown in Fig. 3a. At the level of our experimental resolution, the dissociation line shape is therefore not sensitive to the change in interactions. As illustrated in Fig. 1, the pair size can in principle be obtained from both E_{th} and E_w . However, because the whole spectrum may be subject to shifts due to Hartree terms^{16,18}, in the following we focus only on the width of the spectrum.

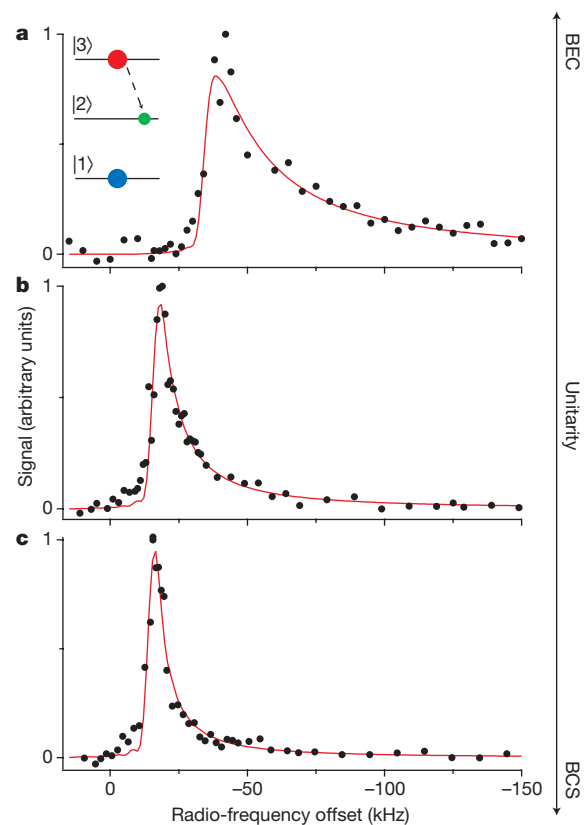


Figure 2 | Radio-frequency dissociation spectra in the BCS-BEC crossover. Below (**a**), at (**b**) and above (**c**) resonance, the spectrum shows the typical asymmetric shape of a pair dissociation spectrum. The signal is proportional to the three-dimensional local response at the centre of the cloud (see Methods). The red traces are fits to the spectra. Because state $|3\rangle$ has a higher energy than state $|2\rangle$ (see the inset in **a**), the dissociation energy is always less than the transition frequency E_0/\hbar for the atomic resonance, and the dissociation spectra therefore appear at energies that are negative relative to E_0 . The inverted frequency axis ensures that the dissociation spectrum is always on the right-hand (or 'positive') side of the origin. The magnetic fields, the local Fermi energies ϵ_F , the temperatures and the interaction strengths $1/k_F a_i$ are 670 G, $\hbar \times (24 \text{ kHz})$, $\sim 0.2T_F$, 0.4 (**a**); 691 G, $\hbar \times (21 \text{ kHz})$, $0.1T_F$, ~ 0 (**b**); 710 G, $\hbar \times (17 \text{ kHz})$, $0.1T_F$, -0.3 (**c**).

At unitarity we determine the full width at half maximum to be $E_w = 0.28(5)\varepsilon_F$, corresponding to a spectroscopic pair size $\xi_w = 2.6(2)/k_F$ (here ε_F is the local Fermi energy and $k_F = \sqrt{2m\varepsilon_F}/\hbar$; the quoted errors are purely statistical). The pairs are therefore smaller than the interparticle spacing l , given by $l = n^{-1/3} = (3\pi^2)^{1/3}/k_F \approx 3.1/k_F$ (where n is the total density), and in units of $1/k_F$ are the smallest reported so far for fermionic superfluids. In high-temperature superconductors, ξ values at optimal doping are in the range $5/k_F$ – $10/k_F$ (ref. 9).

In the simple BEC-to-BCS crossover model, the ratio ξ_{pair}/ξ_w varies from 1 to about 0.4. The fact that ξ_w is smaller than the interparticle spacing suggests the use of the molecular ratio, that is, $\xi \equiv \xi_{\text{pair}} = \xi_w = 2.6/k_F$. Before we compare this result with theoretical predictions, we note that various definitions of the pair size differ by factors on the order of unity¹⁹. With this in mind, we find that the ξ value we observe is larger than the pair size of about $1/k_F$ predicted from a functional integral formulation of the BCS–BEC crossover²⁰. Small fermion pair sizes have been explicitly linked to high critical temperatures (T_c) through the relation $T_c/T_F \approx 0.4/k_F\xi_{\text{pair}}$ (where T_F is the local Fermi temperature), which applies for weak coupling⁹. Inserting the observed ξ value into this relation yields an estimate of $T_c/T_F \approx 0.15$, which lies in the predicted range of values, 0.15 to 0.23 (ref. 21). If we use the asymptotic BCS relation $\Delta = (\hbar^2/\pi m)(k_F/\xi_c) = (\varepsilon_F/\sqrt{2})(1/k_F\xi_{\text{pair}})$, which is valid at weak coupling, and our observed value of ξ at unitarity, we find that $\Delta \approx 0.3\varepsilon_F$, which is smaller than the value of $0.5\varepsilon_F$ predicted by Monte Carlo simulations²².

The strong narrowing of the spectral line in Fig. 2a–c demonstrates that the fermion pair size increases from strong to weak coupling. The decreasing width corresponds to a more than twofold increase of the spectroscopic pair size, from $\xi_w = 1.4(1)k_F$ at 670 G to $\xi_w = 3.6(3)k_F$ at 710 G. A change in the absolute pair size with density at unitarity can in principle be observed by comparing the spectral width in the centre with that in the outer regions of the trapped cloud. As the density decreases, the spectrum shifts to lower energies (Fig. 3b). However, the spectral onset also becomes increasingly softer and

the asymmetry of the pair dissociation peak less pronounced, possibly because of atomic diffusion during the excitation pulse. This prevents a reliable determination of the pair size in the spatial wings, where the density is changing rapidly.

We now consider the effect of final-state interactions in more detail. First we would like to point out that the increase in a_f by a factor of about two between 670 G and 710 G has not affected the spectra in Fig. 3a, within the experimental resolution. This suggests that final-state effects are small for these spectra. Additional information is obtained from the previously introduced bound–bound transitions, which lie outside the range plotted in Fig. 2. On the BEC side of the resonance, the (1, 3) molecule can be transferred also to a more deeply bound (1, 2) molecule (Fig. 4a). The bound–bound peak is still present at unitarity and also on the BCS side at 710 G (Fig. 4b and c), and results from the transition of a many-body bound fermion pair to a (1, 2) molecule. The strong overlap of the pair wavefunction and the molecule in the final state is another indication of the ‘molecular’ character of the fermion pairs in the strongly interacting regime.

The spectra start to change significantly at higher magnetic field. As the magnetic field is increased, the (1, 3) mixture remains in the unitarity-limited regime, with the interaction strength approaching $1/|k_F a_f| \approx 1$ at B_{12} . The final-state interactions, however, change from weak to strong, causing the pair dissociation peak to decrease in weight and the bound–bound peak to become dominant (Fig. 4d–f). This single peak apparently corresponds to a bound–bound transition from many-body bound (1, 3) pairs to the highly correlated final state of an atom in $|2\rangle$ interacting with the paired atoms in $|1\rangle$.

A narrow bound–bound peak is predicted both in the molecular (two-body) and the many-body case, when initial- and final-state interactions are identical or similar. The spectra in Fig. 4 show that bound–bound transitions dominate when $|1/k_F a_i - 1/k_F a_f| \leq 1.5$. In our opinion, a recent theoretical treatment (ref. 23 and S. Basu and E. J. Mueller, personal communication) agrees qualitatively with these results but underestimates the extent of the region in which bound–bound transitions are dominant by a factor of about two. Our observations allow a reinterpretation of the radio-frequency spectra obtained from the (1, 2) superfluid with resonant interactions^{6–8} (see the Supplementary Information for an extended discussion). The spectra have been taken in the regime $|1/k_F a_i - 1/k_F a_f| \leq 1$, where strong bound–bound transitions are expected. Together with the very narrow and symmetric line shape (Supplementary Fig. 3), this suggests that the (1, 2)-to-(1, 3) radio-frequency spectra at 833 G are dominated by such bound–bound transitions and cannot be simply interpreted in terms of a pair dissociation process and a pairing gap^{6–8,24–26}.

In conclusion, we have determined the pair size of resonantly interacting fermions using new superfluid spin mixtures in ^6Li . The (1, 3) mixture is ideally suited to radio-frequency spectroscopy because final-state interactions do not significantly affect the spectra. Our measurements clearly reveal the microscopic structure of the fermion pairs in the strongly interacting regime. The small fermion pair size and high critical temperatures that we observe in our system show an interrelationship similar to the one suggested by the Uemura plot for a wide class of fermionic superfluids⁹. Our results also explain why the rapid-ramp method used to observe fermion pair condensation in the crossover has been successful^{27,28}. The small pair size facilitated the efficient transfer of the many-body bound fermion pairs to more strongly bound molecules, while preserving the momentum distribution of the pairs.

This work presents several opportunities for future research. The microscopic structure of the pairs can now be studied both in the superfluid and normal phase as a function of interaction strength, temperature and spin imbalance between the two components⁷. Increased spectral resolution may reveal interesting deviations of the spectral shape from the generic line shape discussed here.

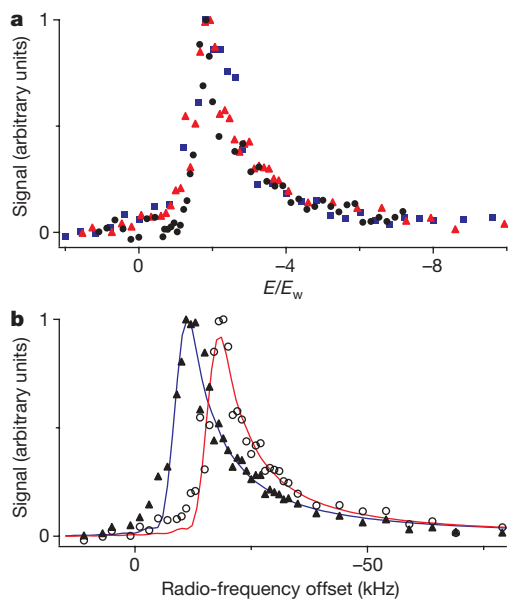


Figure 3 | Comparison of line shapes and density effects. a, Same spectra as in Fig. 2 but with the frequency axis scaled by E_w and shifted so that the spectral onset overlaps with the BEC-side spectrum: black circles, BEC side; red triangles, resonance; blue squares, BCS side. **b**, Density effects at unitarity. The figure shows the tomographically reconstructed spectral response in the centre (open circles, same spectrum as in Fig. 2b; fit by red trace) as well as the lower-density wings (filled triangles; fit by blue trace) of the cloud. In this regime the cloud might be in the normal (as opposed to superfluid) phase.

Furthermore, the predicted universality of a resonantly interacting Fermi mixture can now be tested in ^6Li in three different systems. The lifetimes of all three two-state combinations of the three lowest

hyperfine states in ^6Li are on the order of 10 s in the strongly interacting regime. The three-body decay rates, however, decrease by more than an order of magnitude between 690 and 830 G for a ternary mixture, which may reflect interesting three-body physics. The longer lifetime of 30 ms at 691 G might be sufficient for studies involving all three hyperfine states²⁹, with the potential for experiments on pairing competition in multi-component Fermi gases and spinor Fermi superfluids.

METHODS SUMMARY

Creation of the (1, 3) superfluid. By a method described previously⁵, a spin-polarized sample of ultracold ^6Li in state |1> is obtained in an optical dipole trap after sympathetic cooling with ^{23}Na in a magnetic trap. The equal (1, 3) mixture (that is, one containing the same number of atoms of each component) is prepared at 568 G, close to the zero crossing of a_{13} (the magnetic-field-dependent scattering length between atoms in states |1> and |3>). Here a non-adiabatic Landau–Zener radio-frequency sweep, creating an equal (1, 2) mixture, is followed by an adiabatic Landau–Zener sweep that transfers the atoms in state |2> to state |3>. To induce strong interactions, the magnetic field is adjusted in 100 ms to 730 G and then ramped to values between 660 and 833 G. After evaporative cooling in the optical trap, superfluidity is indirectly established by the observation of fermion pair condensates^{27,28}. Under comparable conditions, quantized vortex lattices—a certain indicator of superfluidity—have been observed in the rotating (1, 2) mixture of ^6Li (ref. 4). $E_F = \hbar^2(v_r^2 v_{ax})^{1/3}(3N)^{1/3}$, with radial trapping frequency $v_r = 140$ Hz, axial trapping frequency $v_{ax} = 22$ Hz and $k_F = \sqrt{2mE_F}/\hbar$. The temperature T/T_F was determined from the shape of the expanded cloud.

Recording the (1, 3) radio-frequency spectra. The radio-frequency dissociation spectra at 670, 691, and 710 G are obtained by applying a 200- μs radio-frequency pulse to the (1, 3) mixture and monitoring the atoms transferred to state |2>. Three-dimensional image reconstruction by means of inverse Abel transformation is used to obtain local radio-frequency spectra⁸. The pulse length is chosen to be shorter than one-quarter the trapping period to minimize atomic diffusion during the excitation pulse. The radio-frequency power is adjusted to transfer fewer than 5% of the total number of atoms. A further reduction of the radio-frequency power affects the signal-to-noise ratio but not the spectral width. The bound–bound spectra and spectra at fields at and above 750 G are not spatially resolved, and are obtained using ~ 1 -ms pulses.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 3 February; accepted 17 June 2008.

- Eagles, D. M. Possible pairing without superconductivity at low carrier concentrations in bulk and thin-film superconducting semiconductors. *Phys. Rev.* **186**, 456–463 (1969).
- Leggett, A. J. in *Modern Trends in the Theory of Condensed Matter* (Proc. 16th Karpacz Winter School Theor. Phys.) (eds Pekalski, A. & Przyslawski, J.) 13–27 (Springer, 1980).
- Nozières, P. & Schmitt-Rink, S. Bose condensation in an attractive fermion gas: from weak to strong coupling superconductivity. *J. Low Temp. Phys.* **59**, 195–211 (1985).
- Zwierlein, M. W., Abo-Shaeer, J. R., Schirotzek, A., Schunck, C. H. & Ketterle, W. Vortices and superfluidity in a strongly interacting Fermi gas. *Nature* **435**, 1047–1051 (2005).
- Ketterle, W. & Zwierlein, M. W. in *Ultra-Cold Fermi Gases* (Proc. Internat. School Phys. ‘Enrico Fermi’, Course 164) (eds Inguscio, M., Ketterle, W. & Salomon, C.) 95–287 (IOS Press, 2008).
- Chin, C. *et al.* Observation of the pairing gap in a strongly interacting Fermi gas. *Science* **305**, 1128–1130 (2004).
- Schunck, C. H., Shin, Y., Schirotzek, A., Zwierlein, M. W. & Ketterle, W. Pairing without superfluidity: The ground state of an imbalanced Fermi mixture. *Science* **316**, 867–870 (2007).
- Shin, Y., Schunck, C. H., Schirotzek, A. & Ketterle, W. Tomographic rf spectroscopy of a trapped Fermi gas at unitarity. *Phys. Rev. Lett.* **99**, 090403 (2007).
- Pistolesi, F. & Strinati, G. C. Evolution from BCS superconductivity to Bose condensation: Role of the parameter $k_F\xi$. *Phys. Rev. B* **49**, 6356–6359 (1994).
- Regal, C. A., Ticknor, C., Bohn, J. L. & Jin, D. S. Creation of ultracold molecules from a Fermi gas of atoms. *Nature* **424**, 47–50 (2003).
- Diener, R. B. & Ho, T.-L. The condition for universality at resonance and direct measurement of pair wavefunctions using rf spectroscopy. Preprint at (<http://xxx.tau.ac.il/abs/cond-mat/0405174>) (2004).
- Yu, Z. & Baym, G. Spin-correlation functions in ultracold paired atomic-fermion systems: Sum rules, self-consistent approximations, and mean fields. *Phys. Rev. A* **73**, 063601 (2006).

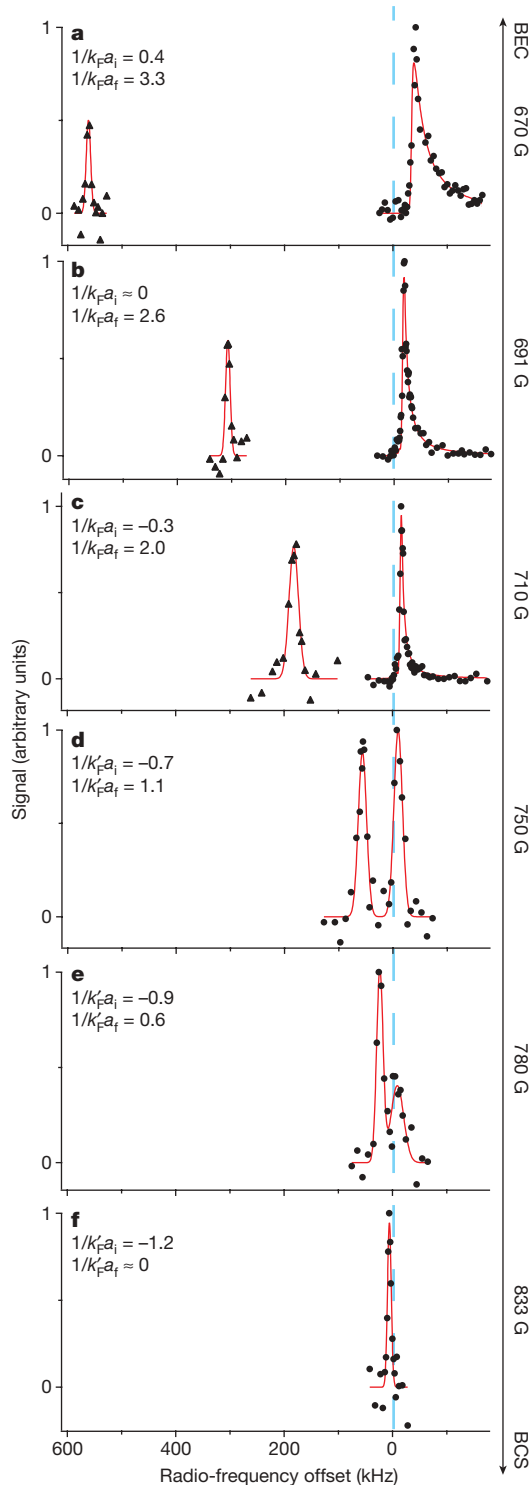


Figure 4 | Effect of final-state interactions on radio-frequency spectroscopy: bound–bound and bound–free spectra in the BCS–BEC crossover of the (1, 3) mixture. Although the initial (1, 3) state is strongly interacting at all fields, the final-state interactions change from weak (a–c) to strong (d–f). See ref. 17 for a plot of the Feshbach resonances. At the higher magnetic fields for $1/k_F a_i \approx -1$, the initial state might be in the normal phase. a–c, Same bound–free spectra and parameters as in Fig. 2. The relative weight of the bound–bound and bound–free peaks could not be determined experimentally (see Methods). d, 750 G, $E_F = \hbar \times (22 \text{ kHz})$, $T/T_F = 0.09$; e, 780 G, $E_F = \hbar \times (23 \text{ kHz})$, $T/T_F = 0.09$; f, 833 G, $E_F = \hbar \times (20 \text{ kHz})$, $T/T_F = 0.06$.

13. Baym, G., Pethick, C. J., Yu, Z. & Zwierlein, M. W. Coherence and clock shifts in ultracold Fermi gases with resonant interactions. *Phys. Rev. Lett.* **99**, 190407 (2007).
14. Punk, M. & Zwerger, W. Theory of rf-spectroscopy of strongly interacting fermions. *Phys. Rev. Lett.* **99**, 170404 (2007).
15. Perali, A., Pieri, P. & Strinati, G. C. Competition between final-state and pairing-gap effects in the radio-frequency spectra of ultracold Fermi atoms. *Phys. Rev. Lett.* **100**, 010402 (2008).
16. Gupta, S. *et al.* Rf spectroscopy of ultracold fermions. *Science* **300**, 1723–1726 (2003).
17. Bartenstein, M. *et al.* Precise determination of ^6Li cold collision parameters by radio-frequency spectroscopy on weakly bound molecules. *Phys. Rev. Lett.* **94**, 103201 (2005).
18. Regal, C. A. & Jin, D. S. Measurement of positive and negative scattering lengths in a Fermi gas of atoms. *Phys. Rev. Lett.* **90**, 230404 (2003).
19. Ortiz, G. & Dukelsky, J. BCS-to-BEC crossover from the exact BCS solution. *Phys. Rev. A* **72**, 043611 (2005).
20. Engelbrecht, J. R., Randeria, M. & Sá de Melo, C. A. R. BCS to Bose crossover: Broken-symmetry state. *Phys. Rev. B* **55**, 15153–15156 (1997).
21. Burovski, E., Prokof'ev, N., Svistunov, B. & Troyer, M. Critical temperature and thermodynamics of attractive fermions at unitarity. *Phys. Rev. Lett.* **96**, 160402 (2006).
22. Carlson, J., Chang, S.-Y., Pandharipande, V. R. & Schmidt, K. E. Superfluid Fermi gases with large scattering length. *Phys. Rev. Lett.* **91**, 050401 (2003).
23. Basu, S. & Mueller, E. J. Final-state effects in the radio frequency spectrum of strongly interacting fermions. Preprint at (<http://xxx.tau.ac.il/abs/0712.1007v2>) (2007).
24. Kinnunen, J., Rodríguez, M. & Törmä, P. Pairing gap and in-gap excitations in trapped fermionic superfluids. *Science* **305**, 1131–1133 (2004).
25. Ohashi, Y. & Griffin, A. Single-particle excitations in a trapped gas of Fermi atoms in the BCS-BEC crossover region. II. Broad Feshbach resonance. *Phys. Rev. A* **72**, 063606 (2005).
26. He, Y., Chen, Q. & Levin, K. Radio-frequency spectroscopy and the pairing gap in trapped Fermi gases. *Phys. Rev. A* **72**, 011602 (2005).
27. Regal, C. A., Greiner, M. & Jin, D. S. Observation of resonance condensation of fermionic atom pairs. *Phys. Rev. Lett.* **92**, 040403 (2004).
28. Zwierlein, M. W. *et al.* Condensation of pairs of fermionic atoms near a Feshbach resonance. *Phys. Rev. Lett.* **92**, 120403 (2004).
29. Honerkamp, C. & Hofstetter, W. Ultracold fermions and the SU(N) Hubbard model. *Phys. Rev. Lett.* **92**, 170403 (2004).
30. Chin, C. & Julienne, P. S. Radio-frequency transitions on weakly bound ultracold molecules. *Phys. Rev. A* **71**, 012713 (2005).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank M. Zwierlein, W. Zwerger, E. Mueller and S. Basu for discussions and A. Keshet for the experiment control software. This work was supported by the NSF and ONR, through a MURI program, and under ARO Award W911NF-07-1-0493 with funds from the DARPA OLE programme.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to C.H.S. (chs@mit.edu).

METHODS

Creation of new superfluid spin mixtures for radio-frequency spectroscopy.

For the well-established (1, 2) mixture, only the $|2\rangle$ -to- $|3\rangle$ transition has been used for radio-frequency spectroscopy. The final state s -wave scattering length a_{13} at B_{12} is large and negative, leading to strong final-state interactions with $1/k_F a_f < -1$ ($a_{13} \approx -3,300a_0$, where a_{ij} is the magnetic-field-dependent scattering length between atoms in states $|i\rangle$ and $|j\rangle$ and a_0 is the Bohr radius). The strength of the final-state interactions can in principle be changed in several ways without affecting the initial state. The density could be lowered to reduce the interaction strength in the final state while the initial state remains resonantly interacting. It is, however, experimentally difficult to decrease the density by a large factor and maintain the same low temperature T/T_F . We might also try spectroscopically to access a different final state. However, in ^6Li there are no other allowed transitions insensitive to the magnetic field. Magnetic field insensitivity is necessary to obtain the required spectral resolution in the kilohertz regime.

Because other mixtures of hyperfine states in ^6Li also exhibit broad Feshbach resonances, we attempt to create resonantly interacting superfluids in new combinations of initial hyperfine states: (1, 3) and (2, 3). The lifetimes of these spin mixtures at resonance exceed 10 s, implying inelastic collision rates smaller than $10^{-14} \text{ cm}^{-3} \text{ s}^{-1}$. For the (2, 3) superfluid the final-state interactions are also large and negative, but the final-state scattering length at B_{13} is either $a_{23} \approx 1,140a_0$ or $a_{12} \approx 1,450a_0$ (depending on the radio-frequency transition employed), and therefore considerably smaller and positive.

Creation of the (2, 3) superfluid. To prepare a (2, 3) superfluid, we follow essentially the same procedure as previously described for the (1, 2) mixture^{4,5}. The only difference is that instead of applying a Landau–Zener transfer that creates an equal (1, 2) mixture, a complete transfer to state $|2\rangle$ is followed by a second sweep, creating an equal (2, 3) mixture. The final magnetic field at the centre of the (2, 3) resonance is $B_{23} \approx 811 \text{ G}$. As in the other spin mixtures, we observe fermion pair condensation after evaporation in the optical trap.

Recording the (1, 3) spectra: stability of the mixture after the radio-frequency pulse. Recording the atoms transferred to state $|2\rangle$ is advantageous because there is no background without a radio-frequency pulse, but it requires that their lifetime with respect to three-body recombination be sufficiently long.

For fields below $\sim 710 \text{ G}$, we find that the lifetime of the $|2\rangle$ atoms after the radio-frequency pulse is short when they form a molecule with a $|1\rangle$ atom as the result of a bound–bound transition. Therefore, in some cases, the bound–bound parts of the spectra are recorded by observing atom number loss in the initial state. After bound–free excitation, the lifetime of atoms in state $|2\rangle$ is 30 ms (determined at 691 G), which is sufficiently long to observe the atoms directly. As a result of the different decay times and recording methods, the relative signal strength of the bound–bound and bound–free parts of the spectrum cannot be determined.

At fields above $\sim 750 \text{ G}$, we find similar and strong losses after both bound–free and bound–bound excitations. Therefore, all data are taken by monitoring losses in the initial state $|3\rangle$ and the spectra reflect the relative strength of bound–bound and bound–free transitions.

Fitting the (1, 3) spectra. The fit to the radio-frequency dissociation spectra in Figs 2, 3 and 4a–c uses a generic, model-independent pair dissociation line shape based on I_m with an additional parameter E_{offset} : $I_{\text{generic}}(E) \propto \sqrt{E - E_{\text{th}}}/(E - E_{\text{offset}})^2$. This line shape provides an excellent fit to simulated radio-frequency dissociation spectra both in the BEC and the BCS limit⁵ (see Supplementary Fig. 2). We used this line shape, convolved with the Fourier transform of the square pulses, as a fitting function, and found good agreement with the experimentally obtained spectra shown in Fig. 2. The generic fit function contains no corrections for final-state interactions. In the BEC limit (where $E_{\text{offset}} = 0$ and $E_{\text{th}} = E_b$), such corrections can be included³⁰ in a multiplicative factor of $1/(E + \hbar^2/m a_f^2 - E_{\text{th}})$. When applied to the dissociation spectra in the crossover, this correction factor changes the fit by only a negligible amount. All bound–bound spectra and the bound–free spectra in Fig. 4d–f have been fit using a Gaussian.

(1, 3) mixture: $|3\rangle$ -to- $|2\rangle$ versus $|1\rangle$ -to- $|2\rangle$ transition. The (1, 3) superfluid gives us the opportunity to record two different (magnetic-field-insensitive) radio-frequency spectra: from state $|3\rangle$ to state $|2\rangle$ (the transition used for all the spectra shown in the paper) and from state $|1\rangle$ to state $|2\rangle$. This allows us to compare radio-frequency spectra of the same system but for somewhat different final-state interactions. The final-state scattering lengths at B_{13} are $a_{23} \approx 1,140a_0$ for the $|1\rangle$ -to- $|2\rangle$ transition and $a_{12} \approx 1,450a_0$ for the $|3\rangle$ -to- $|2\rangle$ transition. Supplementary Fig. 1 shows the spectra at 691 G. Note that the fermion pair size obtained from the spectra agrees for both radio-frequency transitions, within the experimental uncertainty.

LETTERS

Using photoemission spectroscopy to probe a strongly interacting Fermi gas

J. T. Stewart¹, J. P. Gaebler¹ & D. S. Jin¹

Ultracold atomic gases provide model systems in which to study many-body quantum physics. Recent experiments using Fermi gases have demonstrated a phase transition to a superfluid state with strong interparticle interactions^{1–6}. This system provides a realization of the ‘BCS–BEC crossover’⁷ connecting the physics of Bardeen–Cooper–Schrieffer (BCS) superconductivity with that of Bose–Einstein condensates (BECs). Although many aspects of this system have been investigated, it has not yet been possible to measure the single-particle excitation spectrum (a fundamental property directly predicted by many-body theories). Here we use photoemission spectroscopy to directly probe the elementary excitations and energy dispersion in a strongly interacting Fermi gas of ⁴⁰K atoms. In the experiments, a radio-frequency photon ejects an atom from the strongly interacting system by means of a spin-flip transition to a weakly interacting state. We measure the occupied density of single-particle states at the cusp of the BCS–BEC crossover and on the BEC side of the crossover, and compare these results to that for a nearly ideal Fermi gas. We show that, near the critical temperature, the single-particle spectral function is dramatically altered in a way that is consistent with a large pairing gap. Our results probe the many-body physics in a way that could be compared to data for the high-transition-temperature superconductors. As in photoemission spectroscopy for electronic materials, our measurement technique for ultracold atomic gases directly probes low-energy excitations and thus can reveal excitation gaps and/or pseudogaps. Furthermore, this technique can provide an analogue of angle-resolved photoemission spectroscopy for probing anisotropic systems, such as atoms in optical lattice potentials.

As interacting quantum systems with highly tuneable parameters and well-understood two-body interactions, ultracold atom gases provide model systems in which to test condensed matter theories. A challenge for experimenters is to find ways to probe these atom gases that relate directly to condensed matter ideas and enable sensitive searches for new phenomena that can advance our understanding of strongly correlated systems. At a very basic level, the effect of interactions is a modification of the single-particle states. As the amount of interaction is increased, the single-particle eigenstates of the non-interacting case become quasi-particles and phase transitions manifest themselves as qualitative changes in the excitation spectrum, such as the appearance of energy gaps. The single-particle excitation spectrum can be predicted by many-body theory and is a fundamental property of any interacting system.

For electronic systems, photoemission spectroscopy provides a powerful technique to probe the occupied single-particle states⁸. In a typical photoemission spectroscopy experiment, electrons are ejected from a substance through the photoelectric effect (Fig. 1a). The photoelectrons are collected, energy- and momentum-resolved, and counted to give a spectrum of intensity as a function of the

measured kinetic energy, $\epsilon_k = \hbar^2 k^2 / 2m$. Here $\hbar = h/2\pi$, where h is Planck’s constant, and m is the particle mass. By conservation of energy, we can determine the energy of the original single-particle state, E_S , using

$$E_S = \epsilon_k + \varphi - h\nu \quad (1)$$

Here $h\nu$ is the photon energy, φ is the work function of the surface and $E_F - E_S$ is often referred to as the binding energy⁸ (E_F is the Fermi energy).

For ultracold atom gases, radio-frequency spectroscopy has been used to probe a strongly interacting Fermi gas^{2,9–13}. In a typical experiment, a radio-frequency pulse drives atoms into an unoccupied Zeeman spin state, where they are counted to yield a spectrum of counts versus frequency. To date, the radio-frequency outcoupled atoms have not been energy- or momentum-resolved. However, in analogy to electron photoemission spectroscopy, the momentum of the radio-frequency photon is negligible in comparison with the typical momentum of the atoms, and the momenta of the outcoupled atoms are therefore characteristic of the original atom states. Equation (1) applies to photoemission spectroscopy of atom gases, by means of momentum-resolved radio-frequency spectroscopy, if we simply replace the work function φ with the Zeeman energy splitting (Fig. 1b). The extension of photoemission spectroscopy from condensed matter to cold Fermi gases was discussed in ref. 14.

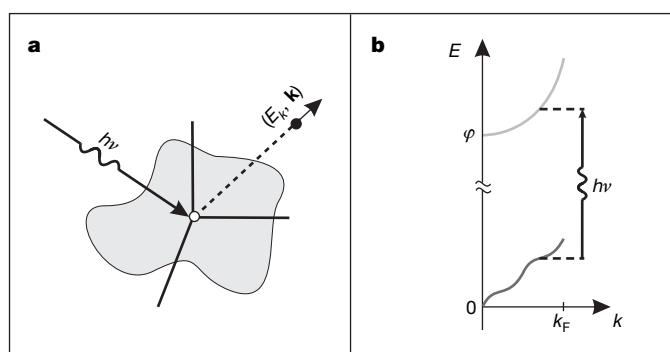


Figure 1 | Photoemission spectroscopy for ultracold atom gases. **a**, In electron photoemission spectroscopy, the energy of electrons emitted from solids, liquids or gases is measured using the photoelectric effect. Using energy conservation, the original energy of the electrons in the substance can be determined. Similarly, in photoemission spectroscopy for atoms, a radio-frequency photon with energy $h\nu$ transfers atoms into a weakly interacting spin state. **b**, The radio-frequency photon drives a vertical transition where the momentum $\hbar k$ is essentially unchanged. By measuring the energy and momentum of the outcoupled atoms (upper curve) we can determine the quasi-particle excitations and their dispersion relation (lower curve). Here φ is the Zeeman energy difference between the two different spin states of the atom.

¹JILA, Quantum Physics Division, National Institute of Standards and Technology and Department of Physics, University of Colorado, Boulder, Colorado 80309-0440, USA.

In this paper, we use photoemission spectroscopy—specifically momentum-resolved radio-frequency spectroscopy—to probe an ultracold gas of fermionic ^{40}K atoms. As in photoemission spectroscopy in solids, this measurement probes the single-particle spectral function, which is directly related to the single-particle Green's function predicted by many-body theories⁸. We use this new technique to probe the Fermi gas near a magnetic-field Fano–Feshbach resonance where strong atom–atom interactions can be tuned to realize a Fermi superfluid in the region of the BCS–BEC crossover^{1–6}.

Our Fermi gas consists of approximately 3×10^5 ^{40}K atoms in a mixture of two spin states. The gas is confined in an optical dipole trap and evaporatively cooled to $T/T_F \approx 0.16$, where T is the temperature, T_F is the Fermi temperature as defined by $T_F = E_F/k_B$, and k_B is Boltzmann's constant. In our photoemission spectroscopy, we apply a radio-frequency pulse that couples atoms in one of the two spin states to an unoccupied, third spin state. There are two essential requirements for determining the excitation spectrum. The first is that the interaction energy be small enough that $\varepsilon_k = \hbar^2 k^2/2m$ holds and the data are not subject to complicated final-state effects^{15–21}. The second requirement is that collisions do not scramble the energy and momentum information carried by the outcoupled atoms. In previous radio-frequency spectroscopy measurements, these requirements were not both satisfied^{2,11–13}. In our ^{40}K gas, however, the interaction energy of the outcoupled atoms is approximately $\hbar \times 640$ Hz, which is much smaller than E_F . Furthermore, the mean free path of the outcoupled atoms is much greater than the size of the gas: $1/\sigma n \approx 6R_F$, where σ is the collision cross-section, n is the average density and R_F is the Fermi radius of the non-interacting gas.

To resolve the kinetic energy ε_k of the radio-frequency outcoupled atoms, we apply a radio-frequency pulse that is short in comparison with the trap period. We then immediately turn off the trap, let the gas ballistically expand and measure the velocity distribution using state-selective time-of-flight absorption imaging (Fig. 2). Assuming a symmetric momentum distribution, we extract the three-dimensional momentum distribution of the outcoupled atoms from the two-dimensional image by performing an inverse Abel transform.

We first consider the case of an ideal Fermi gas. To create a very weakly interacting gas, we adiabatically ramp the magnetic field to the zero crossing of the Feshbach resonance. In Fig. 3a, we plot the intensity, which is proportional to the number of atoms transferred into the third spin state, as a function of the original single-particle energy E_s and wavenumber k . The data are obtained by varying the radio frequency and counting the outcoupled atoms as a function of their momenta. We define zero energy to be the energy of a non-interacting atom at rest in the initial spin state. The intensity map for a non-interacting Fermi gas is expected to show delta function peaks at $E_s = \varepsilon_k$. The white symbols mark the centres of the intensity at each

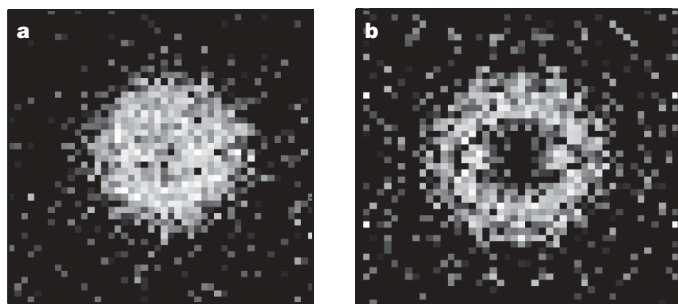


Figure 2 | Extracting the three-dimensional momentum distribution. **a**, A time-of-flight absorption image ($145 \mu\text{m} \times 145 \mu\text{m}$) of atoms that have been transferred into a third spin state is taken after applying a radio-frequency pulse to a Fermi gas on the BEC side of the Feshbach resonance. **b**, After performing quadrant averaging we use an inverse Abel transform to reconstruct the three-dimensional momentum distribution. In this example, a two-dimensional slice at the centre reveals a shell-like structure for the momentum distribution of the outcoupled atoms.

value of k , as determined from Gaussian fits; these show good agreement with the expected dispersion (black line). The root-mean-square width in E_s of the measured spectrum in Fig. 3a is $\hbar \times 2.1$ kHz and is due to an energy resolution that derives from the radio-frequency pulse duration.

To create a strongly interacting Fermi gas, we adiabatically ramp the magnetic field to the peak of the Feshbach resonance, where the s -wave scattering length a diverges and the dimensionless interaction parameter $1/k_F^0 a$ equals zero. Here k_F^0 is the Fermi wavevector that corresponds to the peak density of the original, weakly interacting gas. Previous measurements have shown that after the ramp to $1/k_F^0 a = 0$, our Fermi gas initially with $T/T_F = 0.16$ will be at $(0.9 \pm 0.1) \times T_c$, where T_c is the temperature of the transition to the superfluid state¹. From photoemission spectroscopy of the strongly interacting gas, we extract the intensity map shown in Fig. 3b. The interactions lower the overall energy and flatten the dispersion curve. In addition, the energy width is broadened well beyond our energy resolution.

Interpretation of previous radio-frequency spectroscopy measurements^{2,12} in terms of a pairing gap is a difficult problem that is still unsolved theoretically⁷. The photoemission spectroscopy technique presented here directly measures the occupied density of single-particle states, and is therefore well suited to measuring pairing gaps. In BCS theory the gap vanishes at T_c ; however, in the BCS–BEC crossover a pseudogap due to preformed pairs is predicted to exist above T_c (refs 22–28). The spectral function for a strongly interacting, homogeneous Fermi gas near T_c is predicted to have a gap-like double-peak structure, with peaks following a ‘BCS-like’ dispersion curve

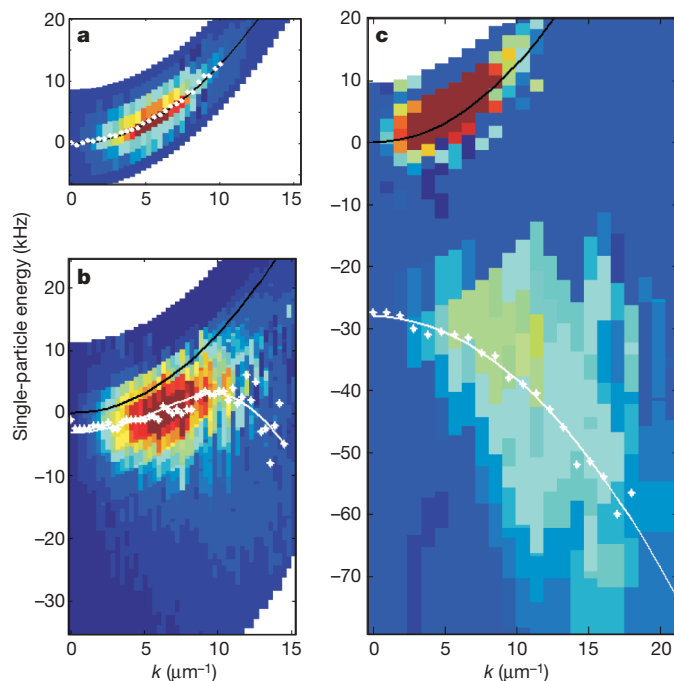


Figure 3 | Single-particle excitation spectra obtained using photoemission spectroscopy of ultracold atoms. Plotted are intensity maps (independently scaled for each plot) of the number of atoms outcoupled to a weakly interacting spin state as a function of the single-particle energy E_s (expressed as frequency) and wavenumber k . Black lines show the expected dispersion curves for an ideal Fermi gas. White symbols mark the centre of each fixed- k energy distribution curve. **a**, Data for a very weakly interacting Fermi gas. The Fermi wavevector k_F^0 is $8.6 \pm 0.3 \mu\text{m}^{-1}$. **b**, Data for a strongly interacting Fermi gas where $1/k_F^0 a = 0$ and $T \approx T_c$. The white line is a fit to the centres to a BCS-like dispersion. **c**, Data for a gas on the BEC side of the resonance where $1/k_F^0 a \approx 1$ and the measured two-body binding energy is $\hbar \times (25 \pm 2)$ kHz. We attribute the upper feature to unpaired atoms and the lower feature to molecules. The white line is a fit to the centres using a quadratic dispersion.

where the BCS gap is replaced by the pseudogap^{22–28}. As a first step in analysing our data, we fit the centres of the intensity at each value of k to this BCS-like dispersion curve, $E_s = \mu - \sqrt{(\epsilon_k - \mu)^2 + \Delta^2}$. Here the fitting parameters are a chemical potential μ and the pseudogap Δ . The best fit, shown as the white curve in Fig. 3b, gives $\mu = h \times (12.6 \pm 0.7 \text{ kHz})$ and $\Delta = h \times (9.5 \pm 0.6 \text{ kHz})$. To facilitate comparison with theory, in Fig. 4 we show measured energy distribution curves for select values of k . We note that in all trapped gas experiments, the density is inhomogeneous and the pairing gap will depend on the local Fermi energy. Therefore, our data should eventually be compared with a theory that includes the effect of the trapping potential through, for example, a local density approximation. Finally, we note that we have performed photoemission spectroscopy for a gas cooled below T_c (initially with $T/T_F = 0.10$) and found that the data are qualitatively similar to those in Fig. 3b.

Far on the BEC side of the resonance, for $1/k_F^0 a \gg 1$, the pairing gap eventually becomes a two-body binding rather than a many-body effect that depends on the local Fermi energy. We measure the excitation spectrum of the Fermi gas at $1/k_F^0 a \approx 1$, where the molecule binding energy measured for a low-density gas is $h \times (25 \pm 2 \text{ kHz})$. We observe two prominent features (Fig. 3c). The first feature is narrow in energy, starts at zero energy and follows the quadratic dispersion expected for free atoms (black line). We attribute this feature to unpaired atoms, which may be out of chemical equilibrium with the pairs. The second feature is very broad in energy, is shifted to lower energy and trends towards lower energy for increasing k . This feature we attribute to atoms in the paired state. An excitation gap separating the two features is evident in the data. We fit the centres of the molecule feature to a quadratic dispersion (white line) with the free parameters being the energy offset and an effective mass m^* . In the BEC limit, where the molecules are tightly bound, we would expect the energy offset to be the molecule binding energy, which equals 2Δ , and the effective mass to be $-m$. This negative effective mass reflects the fact that outcoupling an atom at high k leaves behind an excitation in the form of an unpaired atom. The best fit to the data gives an energy offset of $h \times (28 \text{ kHz})$ and an effective mass of $m^* = -1.25 m$.

The large energy width of the molecule feature seen in Fig. 3c is probably due to centre-of-mass motion of the pairs. For comparison with the data, we have performed a simple Monte Carlo simulation, assuming a thermal distribution for the centre-of-mass motion and using the predicted distribution of relative kinetic energy for radio-frequency dissociation of weakly bound molecules¹⁵. We assume that the pairs are in thermal equilibrium with the unpaired atoms and use the measured temperature of the radio-frequency outcoupled atoms

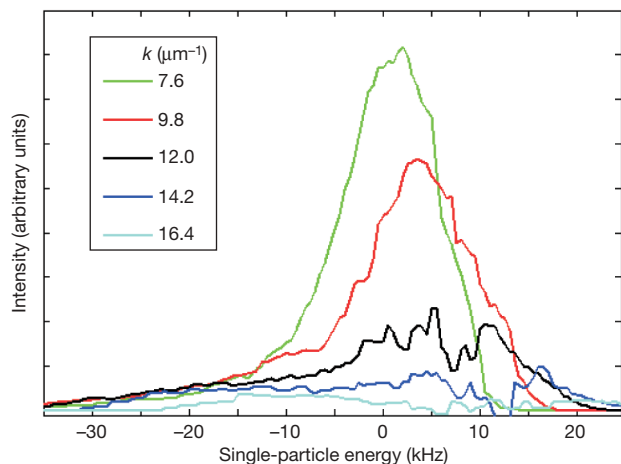


Figure 4 | Energy distribution curves for a strongly interacting Fermi gas. We plot the intensity for selected values of k . Each curve is the average for seven neighbouring values of k in Fig. 3b. The data have been smoothed with a 1.5-kHz-wide filter, and the energy is expressed in terms of frequency.

corresponding to the upper feature in Fig. 3c. Assuming a molecule binding energy of $h \times (25 \text{ kHz})$, we calculate the intensity map shown in Fig. 5a.

The occupied density of states is obtained by summing the data in Fig. 3 over all k (Fig. 5b–d). For the nearly ideal Fermi gas data (Fig. 5b), we find good agreement with the expected density of states for a Fermi gas at $T = 0.18 T_F$ in a harmonic trap (red curve). For the strongly interacting gas (Fig. 5c), the occupied density of states becomes wider in energy and the peak shifts towards lower energies by an amount comparable to the initial E_F of the weakly interacting gas (dashed line). We note that the peak density of the cloud, and therefore the local Fermi energy, increases as we approach the Feshbach resonance²⁹. On the BEC side of the resonance (Fig. 5d), a pairing gap between bound pairs and free atoms is readily apparent. The red curve is the expected density of states determined from the simulation of a thermal distribution of weakly bound molecules (Fig. 5a). The only free parameter in the simulation is an overall scaling factor.

In this work, we have used photoemission spectroscopy, accomplished by momentum-resolving the outcoupled atoms in radio-frequency spectroscopy, to probe the occupied density of single-particle states and energy dispersion through the BCS–BEC crossover. In the future, it may be possible to use spatially resolved photoemission spectroscopy to probe the local pairing gap. Another extension of this work will be to study the BCS–BEC crossover as a function of temperature and/or unbalanced spin population. Photoemission spectroscopy of ultracold atoms is a powerful and conceptually simple probe of strongly correlated atom gases that could be applied to many other atom gas systems. In the work presented here, the atoms

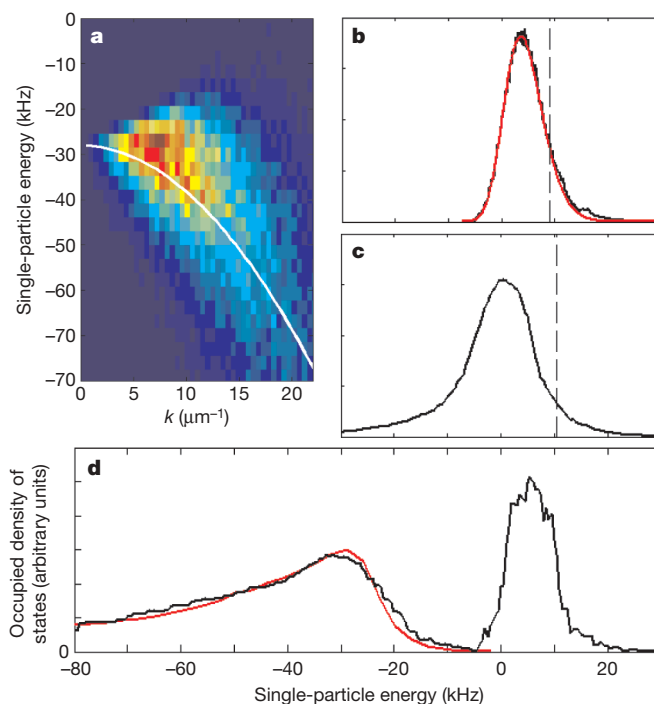


Figure 5 | The occupied density of single-particle states. **a**, A calculated intensity map for a $T = 0.17 \mu\text{K}$ thermal distribution of weakly bound molecules. The white line is the fit to the data shown in Fig. 3c. **b**, The density of states for a weakly interacting Fermi gas (black line) agrees well with a fit (red curve) for a Fermi gas in a harmonic trap. The fit, whose only free parameter is the amplitude, includes our measurement resolution. The dashed black vertical line shows E_F . **c**, The density of states taken at the peak of the Feshbach resonance is shifted to much lower energy. **d**, The density of states on the BEC side of the resonance has two features: a peak due to unpaired atoms and a broader feature due to molecules. The red curve is the expected density of states from the simulation shown in **a**. Energy is expressed in terms of frequency.

interact through isotropic s -wave interactions; therefore, it was not necessary to consider different directions of the outcoupled atoms' momenta. However, like angle-resolved photoemission spectroscopy for solids, this technique could also be applied to non-isotropic systems such as atoms in an optical lattice, low-dimensional systems or higher partial-wave pairing of atoms³⁰.

METHODS SUMMARY

We evaporatively cool an equal mixture of ^{40}K atoms in the $|f, m_f\rangle = |9/2, -7/2\rangle$ and $|f, m_f\rangle = |9/2, -9/2\rangle$ states, where f and m_f give the hyperfine level in the ground-state manifold²⁹. The cylindrically symmetric trap has frequencies $f_r = 234\text{ Hz}$ and $f_z = 19\text{ Hz}$. The imaging beam propagates along the z direction. After evaporation, we adiabatically increase the interaction strength by lowering the magnetic field, at a rate of 0.1 G ms^{-1} , to a value near the Feshbach resonance at $202.10 \pm 0.07\text{ G}$ (ref. 1). For the data in Fig. 3a–c, the magnetic field values are 208.43, 202.10 and 201.51 G, and the total numbers of atoms are 2.6×10^5 , 3.6×10^5 and 2.6×10^5 , respectively.

We apply a radio-frequency pulse with a Gaussian amplitude envelope with a $1/e^2$ width of $240\text{ }\mu\text{s}$, to transfer atoms from the $|9/2, -7/2\rangle$ state to the $|9/2, -5/2\rangle$ state. The radio frequency is approximately 47 MHz. Atoms in the $|9/2, -5/2\rangle$ state have a two-body s -wave scattering length of 130 Bohr radii with the $|9/2, -7/2\rangle$ state and 250 Bohr radii with the $|9/2, -9/2\rangle$ state. Immediately after the radio-frequency pulse, we turn off the optical trap and make a resonant absorption image of the $|9/2, -5/2\rangle$ atoms after 3 to 6.5 ms of expansion. Typically four images are averaged for each radio frequency. For the weakly interacting Fermi gas, the radio-frequency power was chosen to achieve maximum transfer. For the other data, no more than 30% of the atoms were transferred. At the higher radio frequencies, we increased the signal by increasing the radio-frequency power and then scaled the data to give the appropriate intensity.

There is an uncertainty of $\pm 1\text{ kHz}$ in ϕ , which is measured for a weakly interacting gas, and therefore also in the zero of E_s . The measured k has a 5% uncertainty due to the imaging magnification.

Received 17 March; accepted 10 June 2008.

- Regal, C. A., Greiner, M. & Jin, D. S. Observation of resonance condensation of fermionic atom pairs. *Phys. Rev. Lett.* **92**, 040403 (2004).
- Chin, C. *et al.* Observation of the pairing gap in a strongly interacting Fermi gas. *Science* **305**, 1128–1130 (2004).
- Partridge, G. B., Strecker, K. E., Kamar, R. I., Jack, M. W. & Hulet, R. G. Molecular probe of pairing in the BEC-BCS crossover. *Phys. Rev. Lett.* **95**, 020404 (2005).
- Zwierlein, M. W., Abo-Shaeer, J. R., Schirotzek, A., Schunck, C. H. & Ketterle, W. Vortices and superfluidity in a strongly interacting Fermi gas. *Nature* **435**, 1047–1051 (2005).
- Luo, L., Clancy, B., Joseph, J., Kinast, J. & Thomas, J. E. Measurement of the entropy and critical temperature of a strongly interacting Fermi gas. *Phys. Rev. Lett.* **98**, 080402 (2007).
- Tarruell, L. *et al.* in *Ultra-Cold Fermi Gases* (Proc. Internat. School Phys. 'Enrico Fermi', Course 164) (eds Inguscio, M., Ketterle, W. & Salomon, C.) 845–855 (IOS Press, 2008).
- Giorgini, S., Pitaevskii, L. P. & Stringari, S. Theory of ultracold Fermi gases. *Rev. Mod. Phys.* (in the press).
- Damascelli, A. Probing the electronic structure of complex systems by ARPES. *Phys. Scr.* **T109**, 61–74 (2004).
- Regal, C. A. & Jin, D. S. Measurement of positive and negative scattering lengths in a Fermi gas of atoms. *Phys. Rev. Lett.* **90**, 230404 (2003).
- Regal, C. A., Ticknor, C., Bohn, J. L. & Jin, D. S. Creation of ultracold molecules from a Fermi gas of atoms. *Nature* **424**, 47–50 (2003).
- Gupta, S. *et al.* Radiofrequency spectroscopy of ultracold fermions. *Science* **300**, 1723–1726 (2003).
- Schunck, C. H., Shin, Y., Schirotzek, A., Zwierlein, M. W. & Ketterle, W. Pairing without superfluidity: The ground state of an imbalanced Fermi mixture. *Science* **316**, 867–870 (2007).
- Schunck, C. H., Shin, Y., Schirotzek, A. & Ketterle, W. Determination of the fermion pair size in a resonantly interacting superfluid. Preprint at (<http://arxiv.org/abs/0802.0341v1>) (2008).
- Dao, T.-L., Georges, A., Dalibard, J., Salomon, C. & Carusotto, I. Measuring the one-particle excitations of ultracold fermionic atoms by stimulated Raman spectroscopy. *Phys. Rev. Lett.* **98**, 240402 (2007).
- Chin, C. & Julienne, P. S. Radio-frequency transitions on weakly bound ultracold molecules. *Phys. Rev. A* **71**, 012713 (2005).
- Yu, Z. & Baym, G. Spin-correlation functions in ultracold paired atomic-fermion systems: Sum rules, self-consistent approximations, and mean fields. *Phys. Rev. A* **73**, 063601 (2006).
- Punk, M. & Zwerger, W. Theory of rf-spectroscopy of strongly interacting fermions. *Phys. Rev. Lett.* **99**, 170404 (2007).
- Perali, A. & Strinati, G. C. Competition between final-state and pairing gap effects in the radio-frequency spectra of ultracold Fermi atoms. *Phys. Rev. Lett.* **100**, 010402 (2008).
- Basu, S. & Mueller, E. J. Final-state effects in the radio frequency spectrum of strongly interacting fermions. Preprint at (<http://arxiv.org/abs/0712.1007v2>) (2007).
- Veillette, M. *et al.* Radio frequency spectroscopy of a strongly imbalanced Feshbach-resonant Fermi gas. Preprint at (<http://arxiv.org/abs/0803.2517>) (2008).
- He, Y., Chien, C.-C., Chen, Q. & Levin, K. Temperature and final state effects in radio frequency spectroscopy experiments on atomic Fermi gases. Preprint at (<http://arxiv.org/abs/0804.1429v1>) (2008).
- Randeria, M. in *Models and Phenomenology for Conventional and High-Temperature Superconductivity* (Proc. Internat. School Phys. 'Enrico Fermi', Course 136) (eds Iadonisi, G., Schrieffer, J. R. & Chiofalo, M. L.) 53–57 (IOS Press, 1998).
- Janko, B., Maly, J. & Levin, K. Pseudogap effects induced by resonant pair scattering. *Phys. Rev. B* **56**, R11407–R11410 (1997).
- Yanase, Y. & Yamada, K. Theory of pseudogap phenomena in high- T_c cuprates based on the strong coupling superconductivity. *J. Phys. Soc. Jpn* **68**, 2999–3015 (1999).
- Perali, A., Pieri, A., Strinati, G. C. & Castellani, C. Pseudogap and spectral function from superconducting fluctuations to the bosonic limit. *Phys. Rev. B* **66**, 024510 (2002).
- Bruun, G. M. & Baym, G. Bragg spectroscopy of cold atomic Fermi gases. *Phys. Rev. A* **74**, 033623 (2006).
- Bulgac, A., Drut, J. E., Magierski, P. & Wlazlowski, G. Gap and pseudogap of a unitary Fermi gas by quantum Monte Carlo. Preprint at (<http://arxiv.org/abs/0801.1504>) (2008).
- Barnea, N. Superfluid to insulator phase transition in a unitary Fermi gas. Preprint at (<http://arxiv.org/abs/0803.2293>) (2008).
- Stewart, J. T., Gaebler, J. P., Regal, C. A. & Jin, D. S. Potential energy of a ^{40}K Fermi gas in the BCS-BEC crossover. *Phys. Rev. Lett.* **97**, 220406 (2006).
- Gaebler, J. P., Stewart, J. T., Bohn, J. L. & Jin, D. S. p -wave Feshbach molecules. *Phys. Rev. Lett.* **98**, 200403 (2007).

Acknowledgements We acknowledge funding from the US NSF. We thank E. Cornell, D. Dessau and the JILA BEC group for discussions.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to D.S.J. (jin@jilaui.colorado.edu).

LETTERS

A hemispherical electronic eye camera based on compressible silicon optoelectronics

Heung Cho Ko^{1*}, Mark P. Stoykovich^{1*}, Jizhou Song², Viktor Malyarchuk³, Won Mook Choi¹, Chang-Jae Yu¹, Joseph B. Geddes III⁴, Jianliang Xiao⁷, Shuodao Wang⁷, Yonggang Huang^{7,8} & John A. Rogers^{1,2,3,4,5,6}

The human eye is a remarkable imaging device, with many attractive design features^{1,2}. Prominent among these is a hemispherical detector geometry, similar to that found in many other biological systems, that enables a wide field of view and low aberrations with simple, few-component imaging optics^{3–5}. This type of configuration is extremely difficult to achieve using established optoelectronics technologies, owing to the intrinsically planar nature of the patterning, deposition, etching, materials growth and doping methods that exist for fabricating such systems. Here we report strategies that avoid these limitations, and implement them to yield high-performance, hemispherical electronic eye cameras based on single-crystalline silicon. The approach uses wafer-scale optoelectronics formed in unusual, two-dimensionally compressible configurations and elastomeric transfer elements capable of transforming the planar layouts in which the systems are initially fabricated into hemispherical geometries for their final implementation. In a general sense, these methods, taken together with our theoretical analyses of their associated mechanics, provide practical routes for integrating well-developed planar device technologies onto the surfaces of complex curvilinear objects, suitable for diverse applications that cannot be addressed by conventional means.

The ability to implement electronic and optoelectronic systems on non-planar surfaces could be useful not only for hemispherical cameras and other classes of bio-inspired device designs, but also for conformal integration on or in biological systems as monitoring devices, prosthetics and so on. Unfortunately, existing technologies have been developed only for surfaces of rigid, semiconductor wafers or glass plates and, in more recent work, flat plastic sheets and slabs of rubber. None is suitable for the application contemplated here because the mechanical strains needed to accomplish the planar to hemispherical geometrical transformation (for example up to ~40% for compact eye-type cameras) greatly exceed the fracture strains (for example a few percent) of all known electronic materials, particularly the most well-developed inorganics, even in 'wavy' structural layouts⁶.

One strategy to circumvent these limitations involves adapting all of semiconductor processing and lithography for direct use on curvilinear surfaces. Even a single part of this type of multifaceted procedure—for example lithographic patterning on such surfaces^{7–15} (see also the Ball Semiconductor website, <http://www.ballsemi.com/>) with levels of resolution and multilevel registration that begin to approach those that can be easily achieved on planar surfaces—requires solutions to extremely difficult technical challenges. Although some work based on plastic deformation of planar sheets^{16,17}, self-assembly of small chips^{18,19} and folding of elastic membranes^{20,21} have shown some promise, each has drawbacks and all require certain processing steps to

be performed on a hemispherical or curved surface. Partly as a result, none have been used to achieve the type of cameras contemplated here.

In this paper, we introduce a means of producing curvilinear optoelectronics and electronic eye imagers that uses well-established electronic materials and planar processing approaches to create optoelectronic systems on flat, two dimensional surfaces in unusual designs that tolerate compression and stretching to large levels of strain (~50% or more). Conceptually this feature enables planar layouts to be geometrically transformed (that is conformally wrapped) to nearly arbitrary curvilinear shapes. In the example presented here, we use a hemispherical, elastomeric transfer element to accomplish this transformation with an electrically interconnected array of single-crystalline silicon photodiodes and current-blocking p–n junction diodes assembled in a passive matrix layout. The resulting hemispherical focal plane arrays, when combined with imaging optics and hemispherical housings, yield electronic cameras that have overall sizes and shapes comparable to the human eye. Experimental demonstrations and theoretical analyses reveal the key aspects of the optics and mechanics of these systems.

Figure 1 illustrates the main steps in the fabrication. The process begins with the formation of a hemispherical, elastomeric transfer element by casting and curing a liquid prepolymer to poly(dimethylsiloxane) (PDMS) in the gap between opposing convex and concave lenses with matching radii of curvature (~1 cm). A jig specially designed to hold these lenses also provides a raised rim around the perimeter of the resulting piece of PDMS. This transfer element is mounted into a mechanical fixture that provides coordinated radial motion of ten independent paddle arms that each inserts into the rim. Translating the arms of this radial tensioning stage outwards expands the hemisphere. The associated reversible, elastic deformations in the PDMS transform this hemisphere, at sufficiently large tension, into the planar shape of a 'drumhead', such that all points in the PDMS are in biaxial tension. The extent of expansion and the underlying mechanics determine the overall magnitude of this tension.

Separately, conventional planar processing forms a passive matrix focal plane array on a silicon-on-insulator wafer consisting of single-crystalline silicon photodetectors, current-blocking p–n junction diodes and metal (chromium–gold–chromium) interconnects, with films of polymer (polyimide) to support certain regions and to encapsulate the entire system. A critically important design feature is the use of thin, narrow lines to connect nearest-neighbour pixel elements; these structures enable elastic compressibility in the system, as described below. Removing the buried oxide layer of the silicon-on-insulator wafer by etching with concentrated hydrofluoric acid in a manner that leaves the focal plane array supported by polymer posts

¹Department of Materials Science and Engineering, ²Department of Mechanical Science and Engineering, ³Frederick-Seitz Materials Research Laboratory, ⁴Beckman Institute for Advanced Science and Technology, ⁵Department of Electrical and Computer Engineering, ⁶Department of Chemistry, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA. ⁷Department of Mechanical Engineering, ⁸Department of Civil and Environmental Engineering, Northwestern University, Evanston, Illinois 60208, USA.

*These authors contributed equally to this work.

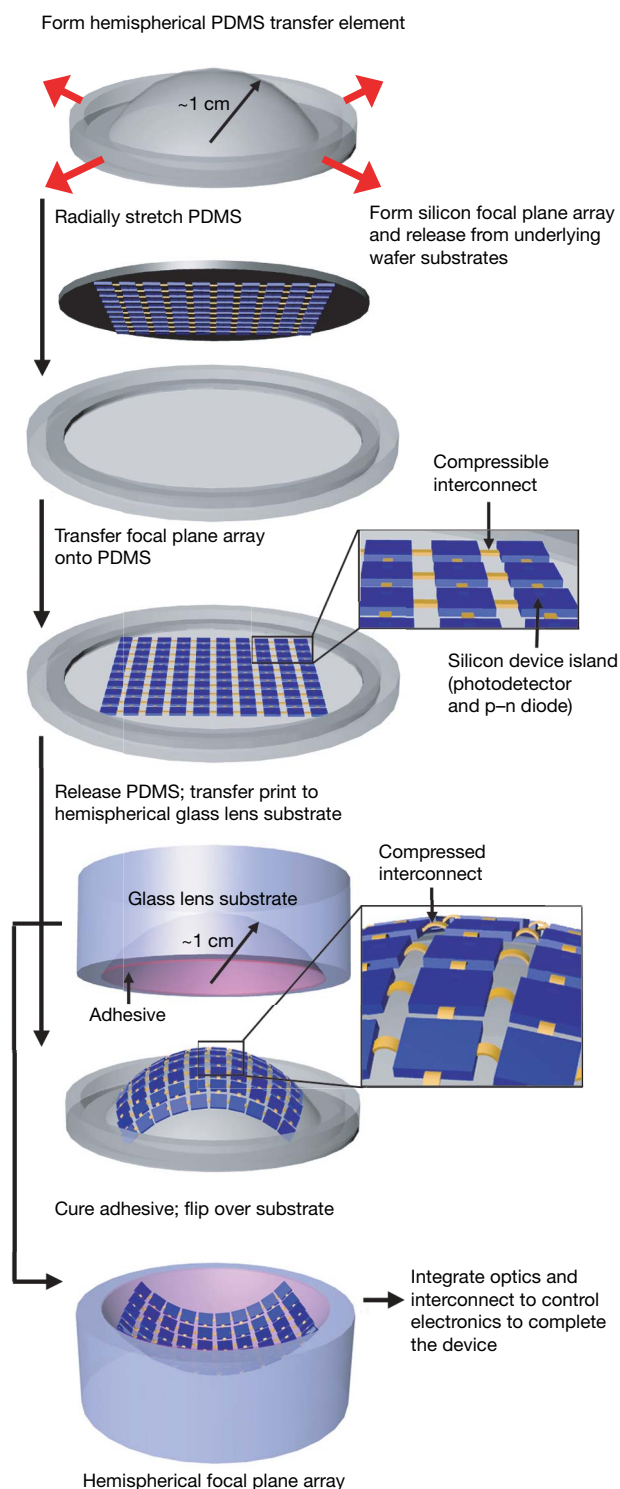


Figure 1 | Illustration of steps for using compressible silicon focal plane arrays and hemispherical, elastomeric transfer elements to fabricate electronic eye cameras. The top frame shows such a transfer element, fabricated in PDMS by casting and curing against an appropriately designed template. Stretching in the radial direction forms a flat drumhead membrane in which all points in the PDMS are in tension. Lifting a prefabricated focal plane array and associated electronics from a source wafer onto the surface of this drumhead, and then allowing the PDMS to relax back to its initial shape, transforms the planar device layout into a hemispherical shape. Transfer printing onto a matching hemispherical glass substrate coated with a thin layer of a photocurable adhesive (pink), adding a hemispherical cap with integrated imaging lens and interfacing to external control electronics (not shown here) completes the camera system.

but otherwise raised above the underlying silicon ‘handle’ wafer completes the device processing. Fabrication of the interconnected pixel arrays on rigid, planar substrates using established processing techniques avoids limitations, for example in registration, that are often encountered in soft electronics. (Fabrication details appear in the Supplementary Information.)

Bringing the transfer element in its tensioned, planar drumhead shape into contact with this wafer and then peeling it away lifts up the focal plane array, leaving it stuck to the soft surface of the elastomer through non-specific van der Waals interactions^{22,23}. Moving the paddle arms of the tensioning stage inwards to their initial positions causes the elastomer to relax back, approximately, to its initial hemispherical shape but with a slightly (~10% for the systems investigated here) larger radius of curvature. In this process, compressive forces act on the focal plane array to bring the pixel elements closer together, with magnitudes that correspond to significant compressive strains (up to 10–20%, depending on the applied tension). The narrow, thin connecting lines accommodate these large strains by delaminating locally from the surface of the elastomer to adopt arc shapes pinned on the ends by the detector pixels (that is, the pixels are not substantially deformed and the interconnect strains are up to ~20–40%), with a mechanics conceptually similar to related responses in stretchable semiconductor ribbons²⁴. This process allows the planar-to-spherical geometrical transformation to be accomplished without creating substantial strains in any of the active components of the focal plane array. The hemispherical, elastomeric transfer element, ‘inked’ with the focal plane array in this manner, enables transfer ‘printing’ onto a hemispherical glass substrate that has a matching radius of curvature and is coated with a thin layer of a photocurable adhesive. Mounting the resulting system on a printed circuit board with bus lines to external control electronics, establishing electrical connections to pinouts located along the perimeter of the detector array, and integrating with a hemispherical cap fitted with a simple imaging lens completes the hemispherical electronic eye camera. (Details appear in the Supplementary Information.)

The fabrication approach summarized in Fig. 1 can be applied to planar electronics and optoelectronics technologies with nearly arbitrary materials classes and devices, provided that they incorporate appropriately configured compressible interconnects. A key advantage of the strategy is that the most labour-intensive part of the process, that is, formation of the pixel arrays themselves, is fully compatible with the capabilities of existing, planar silicon device manufacturing facilities. Figure 2 outlines the designs and processes implemented in constructing the hemispherical cameras described here. Each pixel in the 16-by-16 imaging array supports two devices—a photodetecting diode and a p–n junction blocking diode—monolithically formed in a single piece of single-crystalline silicon ($500 \times 500 \mu\text{m}^2$ in area, $1.2 \mu\text{m}$ thick) with a capping layer of polyimide ($560 \times 560 \mu\text{m}^2$ in area, $1\text{--}1.5 \mu\text{m}$ thick): the first device provides local light detection; the second enables current blocking and enhanced isolation for passive matrix readout. Layers of metal above each of the blocking diodes shield them from light, thereby removing their photoresponses. The layout of these layers, the two devices and the electrical connections are illustrated in the schematic view and optical image in Fig. 2a. The pixel-to-pixel interconnects consist of thin layers of patterned metal ($360 \mu\text{m}$ long, $50 \mu\text{m}$ wide, $3\text{:}150\text{:}3 \text{ nm}$ thick (chromium:gold:chromium)) on thin layers of polyimide ($360 \mu\text{m}$ long, $110 \mu\text{m}$ wide, $1\text{--}1.5 \mu\text{m}$ thick), spin-cast and patterned in conventional ways.

The images in Fig. 2b, c shows a 16-by-16 array of photodetecting diode–blocking diode pixels transferred onto the surface of a hemispherical, elastomeric transfer element, corresponding to the next-to-last diagram in Fig. 1. The arc-shaped interconnections that enable the planar to hemispherical transformation can be seen clearly. The yields associated with the transfer process and the formation of these types of stretchable connections are high; 100% of the pixels and interconnections in the case of the 16-by-16 arrays have been

reproducibly transferred. Greater than 95% yields have also been demonstrated for the transfer of higher density arrays of passive silicon elements ($20 \times 20 \mu\text{m}^2$ in area, 50 nm thick) and nearest-neighbour connections ($20 \mu\text{m}$ long, $4 \mu\text{m}$ wide, 50 nm thick) (see Supplementary Fig. 9).

Significant mechanical deformations in the imaging arrays are generated during the transfer process, specifically during the planar to hemispherical transformation of the elastomeric transfer element. Simple mechanics models, based on plate theory and confirmed using established finite element analysis techniques, have been developed to determine the spatial distributions of pixels during the transfer process, as well as the distributions of stresses and

displacements in the interconnections and silicon pixels (see Supplementary Information). These models indicate that the imaging arrays on the hemispherical surface have very small variations ($\sim 3\%$ maximum to minimum) in the local pitch and a relatively uniform pitch $\sim 10\%$ smaller than the arrays in the planar, as-fabricated geometry. In addition, the mechanics models predict maximum strains of $\sim 0.01\%$ in the silicon pixels and $\sim 0.3\%$ in the metal of the arc-shaped interconnects for the $\sim 20\%$ change in interconnection length ($\sim 10\%$ change in pitch) observed in these systems. Figure 2d is an image of a completed array on a hemispherical glass substrate, corresponding to the last frame in Fig. 1. The high level of engineering control over the fabrication process is evident from the uniformity of the structures that can be transferred to the hemispherical substrate.

Figure 2e shows the current–voltage response of a representative individual pixel in a hemispherical detector array (black, shielded from light; red, exposed to light), addressed by means of row and column electrodes through contact pads at the perimeter of the 16-by-16 array. Similar responses are achieved for individual pixels in planar imaging arrays (see Supplementary Fig. 15). Key features are the strong photoresponse (main frame), the very low reverse bias current (inset), and low crosstalk (inset) between pixels in passive matrix addressing. (Additional details of the device layouts, the electrical properties and the mechanics analysis appear in the Supplementary Information.)

Electrical connections from the row and column contacts at the periphery of the passive matrix array are made to pre-patterned lines on a printed circuit board. The resulting system (Fig. 3a) is interfaced with a computer with specially designed software for acquiring images from the camera. The external electrical connections are formed by evaporating metal over the edge of the concave glass substrate through a flexible shadow mask. Currently, the electrode lines that connect the periphery of the pixel arrays to separate control electronics limit yields and set practical bounds on pixel counts. With unoptimized manual systems, the interconnects from the periphery of the pixel array to the printed circuit board can be registered to an accuracy of $\pm 200 \mu\text{m}$. Integration with a hemispherical cap fitted with a simple, single-element lens that provides the imaging optics completes the camera, as illustrated in Fig. 3b, c.

Figure 3d shows images collected with the hemispherical electronic eye camera of Fig. 3a–c. The optical setup for these results used collimated green light (argon ion laser) to illuminate a printed pattern on a transparency film. The transmitted light passed through a simple plano-convex lens (diameter, 25.4 mm; focal length, 35 mm) to form an image on the hemispherical camera (see Supplementary Information and Supplementary Fig. 16). The left-hand frame of Fig. 3d shows the direct output of the camera for the case of an image

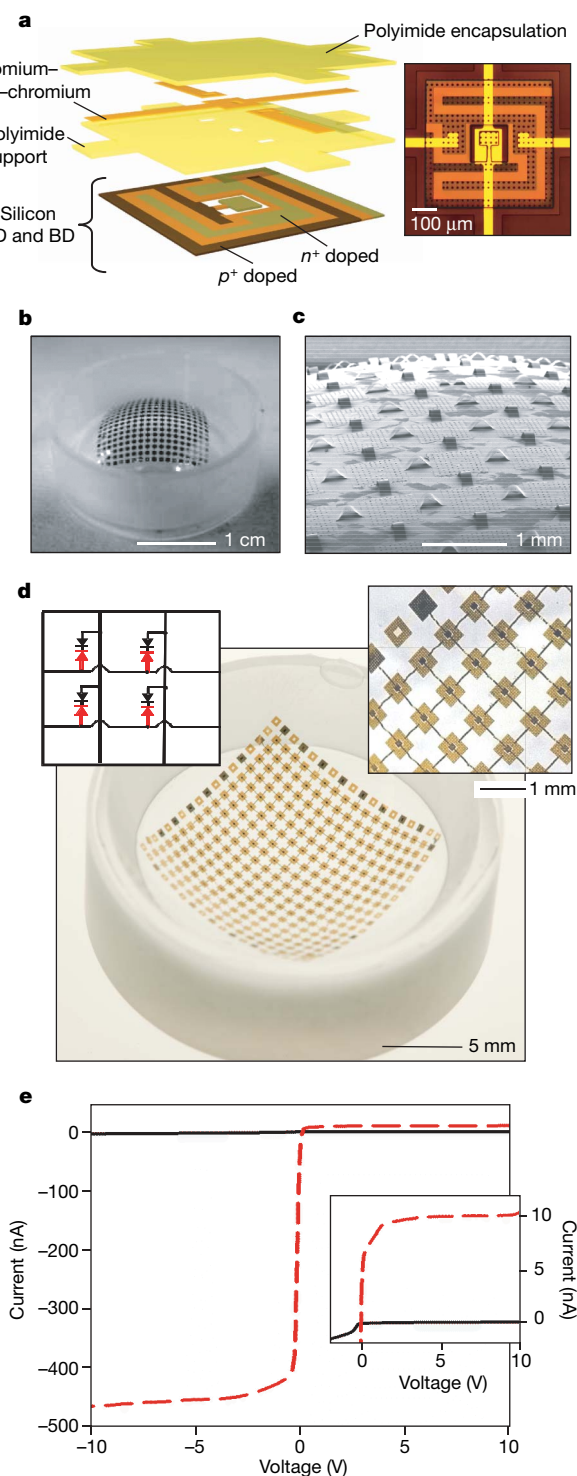


Figure 2 | Design and electrical properties of a hemispherical electronic eye camera based on single-crystalline silicon photodetectors and current-blocking p–n junction diodes in a compressible, passive matrix layout.

a, Exploded schematic of the layout of the silicon, metal and polymer associated with a single unit cell in the array. The blocking diode (BD) is in the centre of the cell; the photodetector (PD) is in a serpentine geometry around the BD. **b**, Photograph of a hemispherical PDMS transfer element with a compressible focal plane array on its surface. **c**, Scanning electron microscope image of a portion of the array in **b**, illustrating the compressible interconnects. **d**, Photograph of the array integrated on a hemispherical glass substrate (main frame), optical micrograph of a part of the array (upper right inset) and circuit diagram showing the BDs (black), PDs (red) and electrode crossovers (arcs) in a 2-by-2 section of the system (upper left inset).

e, Electrical properties of a unit cell. The data were measured by contacting the row and column electrodes that address this position in the hemispherical array, by means of pads at the perimeter of the system. The data (black, shielded from light; red, exposed to light) show a high-contrast response to light exposure. Equally importantly, the reverse bias current and leakage from other pixels in the array are both minimal, as shown in the inset.

of the top two rows of the standard eye chart²⁵. Although the shapes of the letters are clearly resolved, the fine spatial features of the smaller text are not accurately represented, owing to the relatively low numbers of pixels in these cameras. The image quality can be improved by implementing a strategy adapted from biological systems, in which a sequence of images is collected as the camera is eucentrically rotated in the θ and ϕ directions (θ , azimuthal angle in the plane normal to the optical axis; ϕ , polar angle measured from the optical axis) relative to the object. Reconstruction, using pixel positions on the hemispherical surface predicted with mechanics models described in the Supplementary Information, yields high-resolution images. The right-hand frame of Fig. 3d is a picture acquired by rapidly scanning a small range of angles (from -2° to 2° in both θ and ϕ) in 0.4° increments.

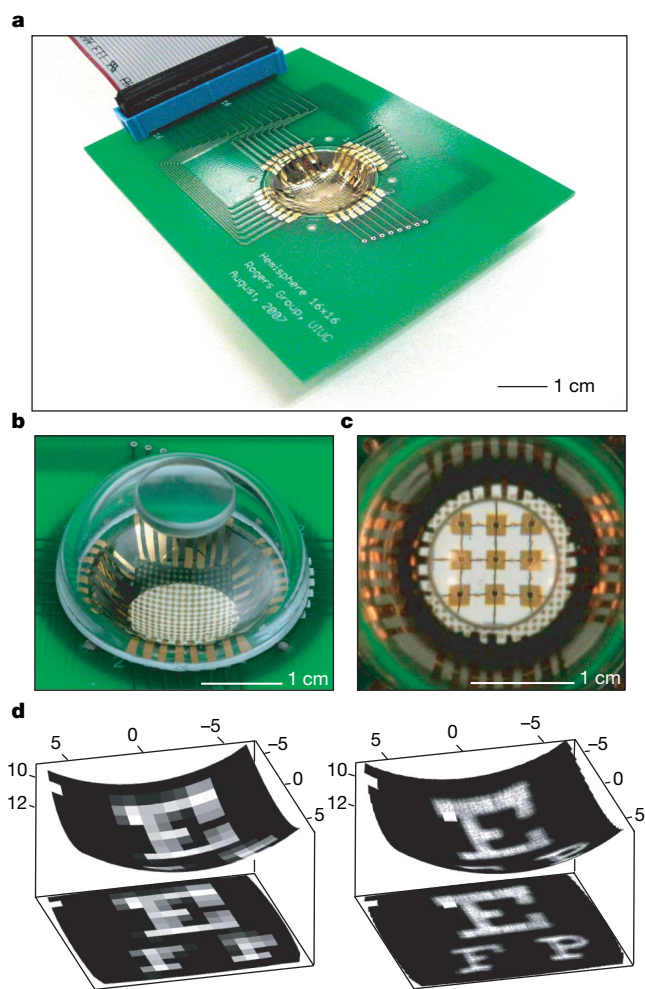


Figure 3 | Photographs of a hemispherical electronic eye camera and representative output images. **a**, Photograph of a hemispherical focal plane array (centre) mounted on a printed circuit board (green), with external connection to a computer (not shown) through a ribbon cable (upper left). **b**, Photograph of the camera after integration with a transparent (for ease of viewing) hemispherical cap with a simple, single-component imaging lens (top). **c**, Close-up photograph of the system in **b**, as viewed directly through the imaging lens. For the parameters used here, this lens magnifies the focal plane array to show a small, 3-by-3 cluster of pixels. **d**, Greyscale images of the first two rows in an eye chart acquired using a hemispherical camera with a 16-by-16 pixel array, as displayed on a hemispherical surface matching the detector surface (top) and projected onto a plane (bottom). The images on the left and right were acquired without scanning and with scanning (from -2° to 2° in the θ and ϕ directions, in 0.4° increments), respectively. The axis scales are in millimetres and are identical in each image.

Even more complex pictures (Fig. 4a, b) can be obtained at high resolution using this simple scanning approach (-2° to 2° in both θ and ϕ , 0.4° increments). Inspection of the images suggests that the stitching errors associated with this process are $<40\ \mu\text{m}$, thereby validating the accuracy of these models. The nearest-neighbour pixels in the hemispherical camera are separated by $\sim 4^\circ$, leading to zero redundancy in generating the tiled picture. These results also demonstrate the high yield of functional pixels, namely $>99\%$ (254 out of 256). Supplementary Figs 18 and 19 show images acquired from each

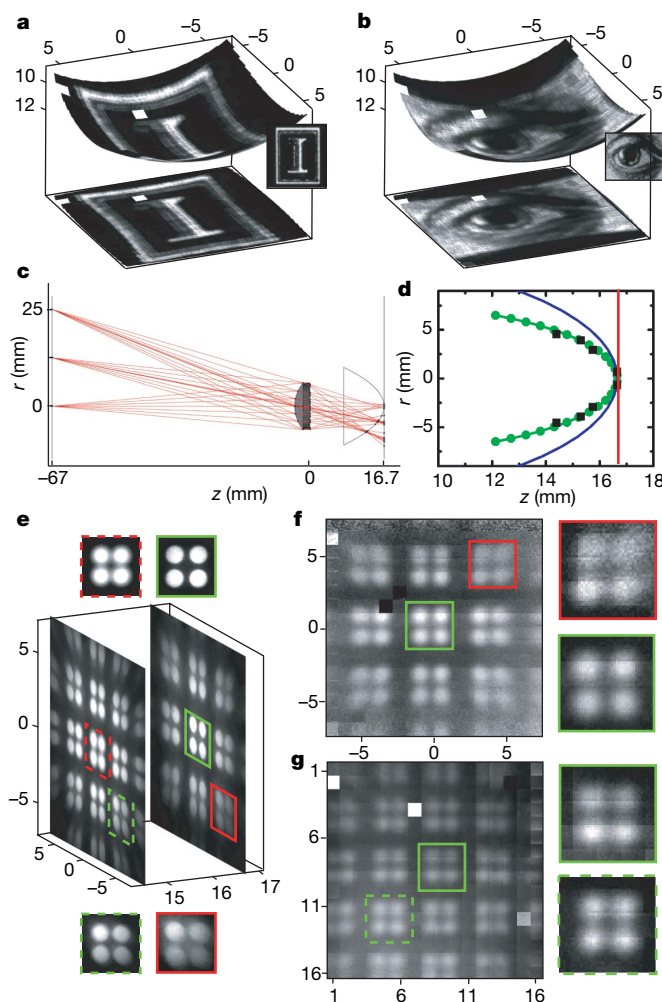


Figure 4 | Enhanced imaging in hemispherical cameras in comparison with planar cameras. High-resolution images of the University of Illinois 'I' logo (**a**) and a drawing of an eye acquired with a hemispherical camera (**b**; insets on the right of **a** and **b** show the original images scanned from the transparency films). **c**, Optics set-up used for imaging and sample ray traces showing a pattern of rays passing through the image and lens onto the detector screens (optimal focal surface and planar camera). Plotted using the Rayica 3.0 package, Optica Software, Champaign, Illinois, USA. **d**, Ray tracing predictions of the curvature of the optimal focal surface (green circles, calculated focal points; green curve, parabolic fit), the detector surface of the hemispherical camera (blue curve), and a planar camera (red curve). **e**, High-resolution photographs of projected images on a planar screen positioned at varying distances from the lens. The left- and right-hand images were acquired at distances $z = 14.40\ \text{mm}$ and $z = 16.65\ \text{mm}$ from the lens, respectively, and demonstrate a shift in optimal focus as a function of detector position. A series of such images was used to estimate the optimal curvilinear focal surface, shown in **d** as black squares. **f**, **g**, High-resolution images acquired with planar (**f**) and hemispherical (**g**) cameras positioned $16.65\ \text{mm}$ from the lens (along the optical axis). All axis scales are in millimetres except for those in **g**, in which pixel number is plotted, and the axes normal to the image planes represent the z direction (optical axis).

pixel when scanned over the entire projected image (from -40° to 40° in both θ and ϕ), further demonstrating the high quality and uniformity of the pixels in the array.

The simple, single-lens system considered here provides a clear example of how curved detectors can improve camera performance. The focusing ability of hemispherical and planar cameras is compared in Fig. 4c–g using fabricated devices, ray tracing software and commercial cameras. An ideal imaging system would perfectly reproduce the image on the detector surface; however, the lens introduces aberrations that degrade the image quality. Complex and expensive optics can reduce the third-order Seidel aberrations for planar detector surfaces, but such aberrations play a significant role in the focusing ability of the simple, single-lens arrangements of interest here^{3,26}. A demonstration of focusing abilities requires non-collimated light sources and a wide aperture for a large field of view; thus, the optical test set-up for Fig. 4c–g uses rear-illumination of a pattern printed on paper with halogen lamps and a high-numerical-aperture plano-convex lens (diameter, 12 mm; focal length, 12 mm). Use of optical filters to limit the incident-light wavelength to ~ 620 – 700 nm minimizes contributions from chromatic aberrations.

Figure 4c shows the optics arrangement and representative ray traces used to calculate the curvilinear image surface (see Supplementary Information). The calculated surface corresponds, to a good approximation, to a paraboloid of revolution (see Fig. 4d) and is much closer in shape to the hemispherical detector than the planar detector. Figure 4e shows images projected on a planar screen (photographic plastic film) obtained with a commercial camera at two different distances z (left, $z = 14.40$ mm; right, $z = 16.65$ mm) between the screen and lens along the optical axis. The position of best focus shifts from the centre to the edge of the image with decreasing z . The image surface estimated using a series of such photographs is similar to that predicted by the ray tracing theory (see Fig. 4d and Supplementary Figure 20). Figure 4f, g compares images acquired with the fabricated planar and hemispherical cameras, respectively. The hemispherical system has a number of advantages, including more uniform focus from the centre to the edge, a wider field of view, more homogeneous intensity throughout the image and reduced geometric distortions. Many of these features are evident in Fig. 4f, g even at the modest levels of resolution associated with these particular devices.

In conclusion, we note that the compressible optoelectronics and elastomeric transfer element strategies introduced here are compatible with high-resolution focal plane arrays, other more advanced materials systems and device designs, and refined substrate shapes (for example parabolic or other aspherical surfaces). Demonstrating these possibilities, exploring applications in other areas, and defining the fundamental limits associated with the materials, layouts and mechanics of the compressible interconnects represent promising topics for future work.

METHODS SUMMARY

Fabrication of hemispherical cameras. A hemispherical, elastomeric transfer element was molded in poly(dimethylsiloxane) (Sylgard 184, Dow Corning) by curing a liquid prepolymer between convex and concave lenses with identical radii of curvature (12.9 mm, CVI Laser Optics). Coordinated radial stretching of the transfer element such that the hemispherical surface could be reversibly deformed to a planar drumhead surface was performed using a custom-designed mechanical stage. A passive matrix focal plane array was fabricated separately on a planar silicon-on-insulator wafer (1.2- μ m-thick silicon layer on a 400-nm-thick silicon dioxide insulator layer, Soitec) using conventional semiconductor processing. The buried oxide layer in the silicon-on-insulator wafer was removed after complete fabrication of the detector array by etching with concentrated hydrofluoric acid and this wafer was contacted to the radially tensioned transfer element in its planar drumhead shape. The focal plane array was lifted onto this planar elastomeric transfer element after peeling away the wafer and remained adhered through intermolecular van der Waals interactions. Relaxing the tension on the planar elastomeric transfer element transformed the focal plane array and transfer element into approximately the original hemispherical surface. The focal plane array was then transfer printed from the hemispherical transfer

element onto a matching concave glass substrate coated with a thin layer of photocurable adhesive (NOA 73, Norland). The hemispherical substrate was mounted in a printed computer board and electrical connections to external pinouts were made from metal evaporated through an elastomeric shadow mask.

Imaging with hemispherical cameras. The hemispherical cameras collected images using two distinct set-ups. Collimated laser light (argon ion laser; Figs 3d and 4a, b) and non-collimated light sources (halogen lamps; Figs 4e–g) illuminated a pattern printed on transparency and paper screens, respectively. The transmitted light then passed through a single plano-convex lens to form an image on the hemispherical camera.

Received 1 February; accepted 20 May 2008.

- Land, M. F. & Nilsson, D.-E. *Animal Eyes* (Oxford Univ. Press, New York, 2002).
- Goldsmith, T. H. Optimization, constraint, and history in the evolution of eyes. *Q. Rev. Biol.* **65**, 281–322 (1990).
- Walther, A. *The Ray and Wave Theory of Lenses* (Cambridge Univ. Press, Cambridge, UK, 1995).
- Swain, P. & Mark, D. Curved CCD detector devices and arrays for multi-spectral astrophysical applications and terrestrial stereo panoramic cameras. *Proc. SPIE* **5499**, 281–301 (2004).
- Grayson, T. Curved focal plane wide field of view telescope design. *Proc. SPIE* **4849**, 269–274 (2002).
- Khang, D. Y., Jiang, H., Huang, Y. & Rogers, J. A. A stretchable form of single crystal silicon for high performance electronics on rubber substrates. *Science* **311**, 208–212 (2006).
- Jackman, R. J., Wilbur, J. L. & Whitesides, G. M. Fabrication of submicrometer features on curved substrates by microcontact printing. *Science* **269**, 664–666 (1995).
- Paul, K. E., Prentiss, M. & Whitesides, G. M. Patterning spherical surfaces at the two-hundred nanometer scale using soft lithography. *Adv. Funct. Mater.* **13**, 259–263 (2003).
- Miller, S. M., Troian, S. M. & Wagner, S. Direct printing of polymer microstructures on flat and spherical surfaces using a letterpress technique. *J. Vac. Sci. Technol. B* **20**, 2320–2327 (2002).
- Childs, W. R. & Nuzzo, R. G. Patterning of thin-film microstructures on non-planar substrate surfaces using decal transfer lithography. *Adv. Mater.* **16**, 1323–1327 (2004).
- Lee, K. J., Fossler, K. A. & Nuzzo, R. G. Fabrication of stable metallic patterns embedded in poly(dimethylsiloxane) and model applications in non-planar electronic and lab-on-a-chip device patterning. *Adv. Funct. Mater.* **15**, 557–566 (2005).
- Lima, O., Tan, L., Goel, A. & Negahban, M. Creating micro- and nanostructures on tubular and spherical surfaces. *J. Vac. Sci. Technol. B* **25**, 2412–2418 (2007).
- Radtke, D. & Zeitner, U. D. Laser-lithography on non-planar surfaces. *Opt. Express* **15**, 1167–1174 (2007).
- Rucheheft, P. & Wolfe, J. C. Optimal strategy for controlling linewidth on spherical focal surface arrays. *J. Vac. Sci. Technol. B* **18**, 3185–3189 (2000).
- Xia, Y. *et al.* Complex optical surfaces formed by replica molding against elastomeric masters. *Science* **273**, 347–349 (1996).
- Hsu, P. I. *et al.* Spherical deformation of compliant substrates with semiconductor device islands. *J. Appl. Phys.* **95**, 705–712 (2004).
- Hsu, P. I. *et al.* Effects of mechanical strain on TFTs on spherical domes. *IEEE Trans. Electron. Dev.* **51**, 371–377 (2004).
- Jacobs, H. O., Tao, A. R., Schwartz, A., Gracias, D. H. & Whitesides, G. M. Fabrication of a cylindrical display by patterned assembly. *Science* **296**, 323–325 (2002).
- Zheng, W., Buhlmann, P. & Jacobs, H. O. Sequential shape-and-solder-directed self-assembly of functional microsystems. *Proc. Natl Acad. Sci. USA* **101**, 12814–12817 (2004).
- Boncheva, M. *et al.* Magnetic self-assembly of three-dimensional surfaces from planar sheets. *Proc. Natl Acad. Sci. USA* **102**, 3924–3929 (2005).
- Boncheva, M. & Whitesides, G. M. Templated self-assembly: Formation of folded structures by relaxation of pre-stressed, planar tapes. The path to ubiquitous and low-cost organic electronic appliances on plastic. *Adv. Mater.* **17**, 553–557 (2005).
- Huang, Y. Y. *et al.* Stamp collapse in soft lithography. *Langmuir* **21**, 8058–8068 (2005).
- Zhou, W. *et al.* Mechanism for stamp collapse in soft lithography. *Appl. Phys. Lett.* **87**, 251925 (2005).
- Sun, Y. *et al.* Controlled buckling of semiconductor nanoribbons for stretchable electronics. *Nature Nanotechnol.* **1**, 201–207 (2006).
- Begbie, G. H. *Seeing and the Eye* 92–93 (Natural History Press, Garden City, New York, 1969).
- Born, M. & Wolf, E. *Principles of Optics* 7th edn (Cambridge Univ. Press, New York, 1999).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank T. Banks, K. Colravy, and J. A. N. T. Soares for help using facilities at the Frederick Seitz Materials Research Laboratory. The materials and optics aspects were developed in work supported by the US Department of Energy, Division of Materials Sciences under Award No. DE-FG02-07ER46471,

through the Materials Research Laboratory and Center for Microanalysis of Materials (DE-FG02-07ER46453) at the University of Illinois at Urbana-Champaign. The processing approaches and the mechanics were developed in work supported by the National Science Foundation under grant DMI-0328162. C.-J.Y. acknowledges financial support from the Korea Research Foundation (grant KRF-2005-214-D00329) funded by the Korean Government (MOEHRD). J.B.G. acknowledges support from a Beckman postdoctoral fellowship.

Author Contributions H.C.K., M.P.S., V.M. and J.A.R. designed the experiments. H.C.K., M.P.S., J.S., V.M., W.M.C., C.-J.Y., J.B.G., J.X., S.W., Y.H. and J.A.R. performed the experiments and analysis. H.C.K., M.P.S., J.S., Y.H. and J.A.R. wrote the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to J.A.R. (jrogers@uiuc.edu) and Y.H. (y-huang@northwestern.edu).

LETTERS

Rapid change in drift of the Australian plate records collision with Ontong Java plateau

Kurt M. Knesel¹, Benjamin E. Cohen¹, Paulo M. Vasconcelos¹ & David S. Thiede¹

The subduction of oceanic plateaux, which contain extraordinarily thick basaltic crust and are the marine counterparts of continental flood-basalt provinces, is an important factor in many current models of plate motion^{1–4} and provides a potential mechanism for triggering plate reorganization⁵. To evaluate such models, it is essential to decipher the history of the collision between the largest and thickest of the world's oceanic plateaux, the Ontong Java plateau, and the Australian plate, but this has been hindered by poor constraints for the arrival of the plateau at the Melanesian trench. Here we present ⁴⁰Ar–³⁹Ar geochronological data on hotspot volcanoes in eastern Australia that reveal a strong link between collision of the Greenland-sized Ontong Java plateau with the Melanesian arc and motion of the Australian plate. The new ages define a short-lived period of reduced northward plate motion between 26 and 23 Myr ago, coincident with an eastward offset in the contemporaneous tracks of seamount chains in the Tasman Sea east of Australia. These features record a brief westward deflection of the Australian plate as the plateau entered and choked the Melanesian trench 26 Myr ago. From 23 Myr ago, Australia returned to a rapid northerly trajectory at roughly the same time that southwest-directed subduction began along the Trobriand trough⁶. The timing and brevity of this collisional event correlate well with offsets in hotspot seamount tracks on the Pacific plate, including the archetypal Hawaiian chain⁷, and thus provide strong evidence that immense oceanic plateaux, like the Ontong Java, can contribute to initiating rapid change in plate boundaries and motions on a global scale.

The Ontong Java plateau (OJP) represents the product of cataclysmic outpouring of magma thought to be associated with the plume-head stage of a mantle hotspot⁸. Occupying an area of 1.5×10^6 km² and with a thickness of 30–35 km, the oceanic plateau is estimated to have a crustal volume of the order of 5×10^7 km³, making it the largest of the world's large igneous provinces⁹. Although the hotspot source remains elusive, widespread emplacement of the OJP began at about 122 Myr ago in the south Pacific⁸. From this time on, the OJP has drifted northward with the Pacific plate⁷. However, the onset of fast northerly motion of Australia, at roughly the same time that the Pacific plate changed to a more westerly trajectory¹⁰, eventually set the massive oceanic plateau on a collision course with the north-eastern margin of the Australian plate. Before the collision, Australia–Pacific convergence was accommodated by southwest-dipping subduction at the Melanesian (Manus–North Solomon–Vitiāz) trench. But, in what has been called a ‘soft-docking’ event (that is, without significant deformation within the plateau)¹¹, the anomalously thick lithosphere of the OJP jammed the trench and ultimately led to a reversal of subduction polarity¹². Thus, although minor southwest-directed subduction has proceeded intermittently and locally in the Solomon Islands¹¹, the Australian plate now subducts to the northeast below the Pacific along the New Britain/San

Cristobal/New Hebrides arc system. Most palaeogeographic reconstructions^{2,3,6,7} place the OJP in the vicinity of the Melanesian trench in the early Miocene epoch (Fig. 1), but the exact timing of the arc–plateau collision has proved difficult to constrain and remains contentious^{11,13}. The main limitation is the scarcity of reliable age data for both subduction-related volcanism, which provides only a fragmented history of plate convergence in the southwest Pacific, and the basement rocks of the OJP, which are largely submerged and buried beneath marine sediments. However, the geometry and age distribution of hotspot-produced volcanic chains provide an alternative forensic tool in the search for this evidence.

Since the early days of plate-tectonic theory, it has been recognized that many intraplate volcanoes are best explained as the product of hotspots: long-lived, relatively stationary points of volcanic activity driven by plumes of deep upwelling mantle¹⁴. Although the number¹⁵, the fixity¹⁶ and even the existence¹⁷ of plume-fed hotspots have all been scrutinized over the ensuing three decades, a consensus holds that a variety of hotspots do exist^{18–20}, and, even when they seem to drift, the volcanic chains that they produce can still help us to decipher the motion and interaction of Earth's tectonic plates^{3,21}. The history of the Ontong Java plateau is no exception. The predicted hotspot trail and seafloor flow-line for the OJP⁷ indicate that, as the massive plateau entered and choked the Melanesian trench, the Australian plate displaced the Pacific plate northward. It has been suggested that this event is linked to southward offsets in the Cobb, Hawaiian and Louisville chains in the north, central and south Pacific at roughly 25 Myr ago⁷ and it may also be recorded in offsets in the seamount tracks in the Tasman Sea (Fig. 2). But, unlike the well-dated Hawaiian–Emperor chain, volcanic samples and age data are sparse for the seamounts east of Australia²², and they do not allow independent testing of this hypothesis. Contemporaneous and more accessible hotspot activity on the eastern margin of the Australian continent, however, may provide evidence that links changes in motion of the Australian plate with the proposed early Miocene collision with the OJP.

Previous geochronological investigations²³ reveal that central volcanoes in eastern Australia become progressively younger to the south, with an average migration rate of 65^{+3}_{-3} km Myr^{–1} for about the past 35 million years. Although consistent with the hypothesis that they, like the seamount tracks to the east, formed as a result of the northward movement of the Australian plate over a stationary mantle hotspot²³, the K–Ar age data do not resolve any variation in age progression during the time period proposed for the early arrival of the OJP. However, recent studies demonstrate that the increased reliability and precision of the ⁴⁰Ar–³⁹Ar method can resolve changes in volcanic migration rates where none were previously apparent^{10,24}. Improved resolution comes in part from the reduced sample size requirements of modern extraction-line/mass-spectrometer systems, combined with incremental heating and isotope-correlation diagram

¹Earth Sciences, The University of Queensland, St Lucia, Queensland, 4072, Australia.

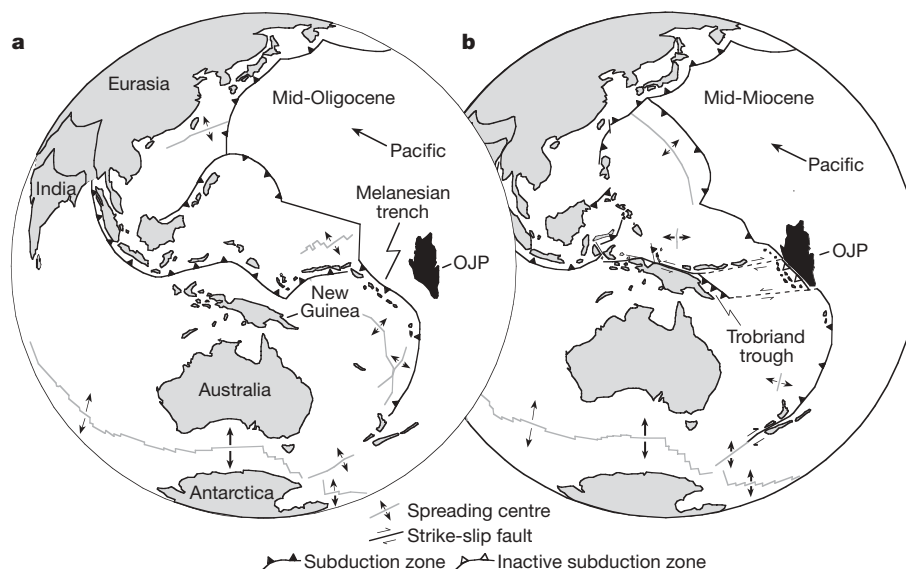


Figure 1 | Simplified reconstructions of the western Pacific. **a**, Mid-Oligocene epoch. **b**, Mid-Miocene epoch. Reconstructions are based on ref. 2, modified to include postulated transform boundaries (dashed lines) connecting the Trobriand and Melanesian subduction zones⁶. The ages of the reconstructions are approximate, and thus, when coupled with the uncertain extent of the OJP before the collision, and the inherent

uncertainty and inevitable oversimplification in any plate model of such a large region^{2,3,6}, they serve only as a rough guide to the timing of OJP encroachment, jamming of the Melanesian subduction zone in the Solomon Islands/Vanuatu/Fiji region, and initiation of southwest-dipping (Maramuni–Trobriand) subduction in eastern New Guinea.

interpretation, which permit the identification of excess argon, argon loss and other sample-related problems that may render some previous K–Ar geochronological results imprecise or inaccurate. We have used this approach, and the striking result is that our 18 new ^{40}Ar – ^{39}Ar ages, when combined with 25 published ^{40}Ar – ^{39}Ar ages^{25,26}, allow resolution of a short-lived period of substantial change in the velocity of the Australian plate between 26 and 23 Myr ago.

All the ^{40}Ar – ^{39}Ar results are presented in Supplementary Tables 1–3 and Supplementary Figures 1–6. The ^{40}Ar – ^{39}Ar data are from petrographically well-characterized samples from nine Cenozoic volcanic centres, which span a north–south distance of roughly 800 km in eastern Australia (Fig. 2). These compositionally bimodal volcanoes contain silicic plugs and/or lava flows, intercalated with and/or surrounded by variably eroded shields of mafic lavas, and are classified as central volcanoes²³. The ^{40}Ar – ^{39}Ar ages confirm that the main period of activity at individual volcanoes is of the order of 3–5 million years²³ and reveal that the silicic products are emplaced over 1 million years or less towards the end of this lifespan (see Supplementary Table 2). The silicic rocks are also more potassium-rich, contain more potassium-rich phenocrysts and are less subject to post-extrusion weathering, yielding more accurate and precise results than mafic rocks. Therefore the age of the silicic rocks from each volcano is the most reliable way of tracking the migration of magmatism through time.

The silicic plugs and lava flows range in age from 31 to 16 Myr and define three migration rates from north to south (Fig. 3). Between Fraser and Flinders they show a linear relationship between age and latitude from 31 to 26 Myr, which yields a migration rate of 71^{+7}_{-4} km Myr^{−1}. The southward migration of volcanism then slowed markedly to 26^{+5}_{-3} km Myr^{−1} before recovering by 23 Myr and remaining steady at 61^{+3}_{-3} km Myr^{−1} until at least 16 Myr. Importantly, the abrupt changes in migration rate match well with the locations of the s-shaped bends in the Tasmanid and Lord Howe seamount chains (Fig. 2a) and with detailed palaeomagnetic investigations of the Tweed and Main Range volcanoes²⁷, which show evidence for a westward plate excursion during this time^{27,28}. Collectively, these results suggest that the Australian plate experienced a period of altered motion, relative to the fixed hotspot frame of reference, from a fast northeasterly trajectory before 26 Myr, changing briefly to a slow northwesterly path, before returning to fast northeasterly

motion by 23 Myr. We also note that, although such changes in absolute plate motion need not necessarily indicate commensurate changes in relative plate motion¹, Australia–Antarctica spreading rates also slowed around 26 Myr, but the available seafloor magnetic record does not seem to fully resolve the short-lived excursion identified here (see Supplementary Fig. 7 for discussion of magnetic data).

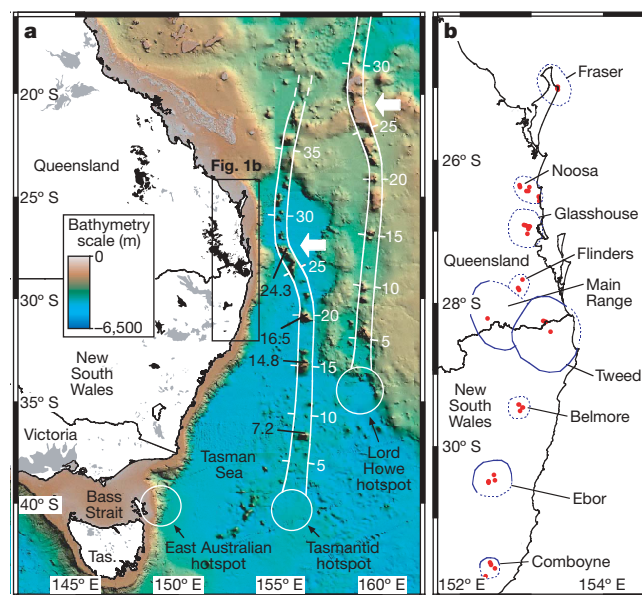


Figure 2 | Distribution of Cenozoic volcanism on the Australian plate. **a**, White arrows show location of s-shaped bends in the Tasmanid and Lord Howe seamount tracks. Hotspot-derived central volcanoes are shaded black; non-hotspot mafic lava fields are grey, and seamount tracks are outlined in white. White circles show predicted present-day hotspot locations. The oldest total-fusion ^{40}Ar – ^{39}Ar ages for Tasmanid seamounts²² are shown in black to the left of the chain; calculated ages are shown in white to the right of the two tracks at 5-Myr intervals (see online Methods for details). **b**, Locations of silicic rocks (shown in red) from the central volcanoes (outlined in blue and dashed where approximate) sampled for ^{40}Ar – ^{39}Ar geochronology.

Our data demonstrate that northward displacement of the Pacific plate at around 25 Myr (ref. 7) was accompanied by a reduction in the northward velocity and a westward deflection of the Australian plate. Consistent with the highly oblique convergence of the OJP from the east, these abrupt and coupled changes in plate motion provide strong evidence for first contact between the plateau and the Melanesian arc at 26 Myr. However, the rapid recovery of fast northward motion of the Australian plate by 23 Myr (Fig. 3), at roughly the same time that the Pacific resumed its westerly motion⁷, also indicates that the initial docking of the OJP was a short-lived event, lasting no more than 3 Myr. This underscores the rapid speed of collision-driven tectonic change and is entirely consistent with the reorganization of plate boundaries in the southwest Pacific around this time². With the jamming of Melanesian subduction, Pacific–Australia convergence was first accommodated by the onset of south-west-dipping subduction of the southern edge of the Solomon Sea along the Trobriand trough (Fig. 1b), forming the Maramuni arc by 20 Myr (ref. 6). But, with the ensuing polarity reversal along the Melanesian arc, subduction in eastern New Guinea ceased by about 10 Myr, and from this time on, convergence was accommodated by subduction of the Australian plate to the northeast^{2,6}. Yet uplift and accretion of the upper part of the OJP in the Solomon Islands did not begin until much later, at about 4 Myr, following subduction of the young and buoyant lithosphere and active spreading system of the Woodlark basin¹¹. Thus, our data support the suggestion that, although some tectonic changes have clearly influenced the structural evolution and deformation of the OJP, at other times the immense plateau, with its greatly thickened crust and deep mantle root, has contributed to initiating rapid change in the absolute motion and boundaries of tectonic plates^{2,7}.

Of course, our evidence for a brief period of altered plate motion between 26–23 Myr rests on the assumption that central volcanoes in eastern Australia are generated over a relatively stationary hotspot. Agreement between relative plate motions from seafloor magnetic anomalies and those from some Indo-Atlantic hotspot traces, including the Tasmanid seamounts east of Australia, support this assumption, indicating little motion among hotspots in this hemisphere during most of the Cenozoic²⁹. Slower plate motion between 26 and 23 Myr ago is also consistent with abrupt changes in the volumetric and compositional characteristics of volcanism in eastern

Australia. For example, the largest central volcano, the Tweed shield, was emplaced during this interval and is estimated to have had a pre-erosion volume of 3,000–4,000 km³, whereas volume estimates for central volcanoes constructed before and after 26–23 Myr range from 800 to less than 100 km³ (ref. 23). The frequency of silicic plugs and lava flows emplaced during this brief time window is also anomalously high, and these rocks comprise a significant proportion of rhyolites with geochemical signatures consistent with a crustal origin³⁰. In contrast, the majority of silicic rocks emplaced outside the 26–23 Myr period are the products of differentiation of mantle-derived magmas, with variable but minor crustal contributions³⁰. Thus, the slower plate motion, and hence longer time spent over the hotspot between 26 and 23 Myr, seems to have promoted construction of larger, more voluminous volcanoes and led to greater melting of the crust.

As expected, the rate and direction changes for the East Australian, Tasmanid and Lord Howe chains occur at about the same time, suggesting that the underlying cause for the shift in the focus of volcanism must be the same. If the shift is not caused by a change in plate motion, but is instead generated by plume movement, then all three plumes must have moved abruptly, concurrently, and in the same direction. This hypothesis is testable by palaeomagnetic and geochronological study of all three tracks, if suitable samples could be obtained by ocean drilling for the seamount chains.

METHODS SUMMARY

Based on the degree of weathering or alteration, phenocryst mineralogy and matrix grain size, we selected samples for either single-crystal (anorthoclase, biotite, hornblende) or groundmass ⁴⁰Ar–³⁹Ar laser incremental-heating analysis. When sanidine occurred as phenocrysts, we analysed between eight and ten aliquots of single crystals by the total-fusion method. The samples were loaded into aluminium disks, along with the neutron fluence monitor Fish Canyon Sanidine, and irradiated for 14 h in the CLICIT facility, a TRIGA-type reactor (Oregon State University). We analysed argon isotopes on a fully automated MAP215-50 instrument at the University of Queensland Argon Geochronology in Earth Sciences (UQ-AGES) laboratory. Samples were incrementally heated using a continuous-wave Ar-ion laser with a 2-mm-wide defocused beam, usually in duplicate, or totally fused under a focused beam. We corrected the data for mass discrimination, atmospheric contamination and nucleogenic interferences using MassSpec Version 7.527 for Macintosh Operating System 10. All replicate incremental-heating results are indistinguishable, and total-fusion analyses of sanidine are highly concordant. Unless indicated otherwise, errors are reported at the 95% confidence interval (2σ) and include errors in the irradiation correction factors and the error in the neutron fluence parameter *J*; they do not include the uncertainty in the potassium decay constants.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 26 February; accepted 3 June 2008.

- Wessel, P. & Kroenke, L. W. Ontong Java Plateau and late Neogene changes in Pacific plate motion. *J. Geophys. Res.* **105**, 28255–28277 (2000).
- Hall, R. Cenozoic geological and plate tectonic evolution of SE Asia and the SW Pacific: Computer-based reconstructions, model and animations. *J. Asian Earth Sci.* **20**, 353–431 (2002).
- Gaina, C. & Müller, D. Cenozoic tectonic and depth/age evolution of the Indonesian gateway and associated back-arc basins. *Earth Sci. Rev.* **83**, 177–203 (2007).
- Wessel, P. & Kroenke, L. W. Reconciling late Neogene Pacific absolute and relative plate motion changes. *Geochim. Geophys. Geosyst.* **8**, doi:10.1029/2007GC001636 (2007).
- Cloos, M. Lithospheric buoyancy and collisional orogenesis: Subduction of oceanic plateaus, continental margins, island arcs, spreading ridges, and seamounts. *Geol. Soc. Am. Bull.* **105**, 715–737 (1993).
- Quarles van Ufford, A. & Cloos, M. Cenozoic tectonics of New Guinea. *Am. Assoc. Petrol. Geol. Bull.* **89**, 119–140 (2005).
- Kroenke, L. W., Wessel, P. & Sterling, A. In *Origin and Evolution of the Ontong Java Plateau* (eds Fitton, J. G., Mahoney, J. J., Wallace, P. J. & Saunders, A. D.) Geol. Soc. London Spec. Publ., **229**, 9–20 (Geol. Soc. Publishing House, Bath, 2004).
- Neal, C. R., Mahoney, J. J., Kroenke, L. W., Duncan, R. A. & Petterson, M. G. in *Large Igneous Provinces: Continental, Oceanic, and Planetary Flood Volcanism* (eds

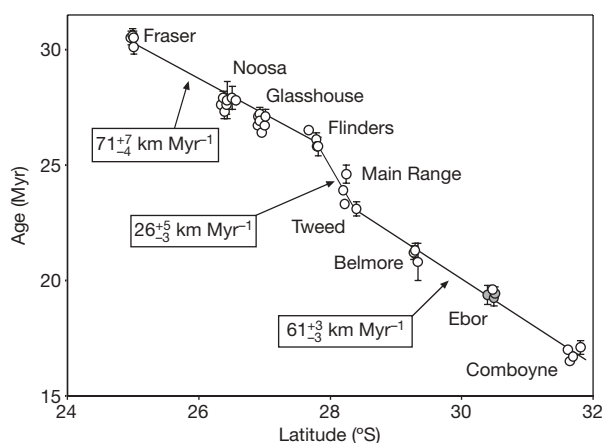


Figure 3 | Ages obtained by ⁴⁰Ar–³⁹Ar dating for silicic volcanic rocks versus latitude, showing abrupt change in volcanic migration rate. Ages for the Fraser, Noosa, Glasshouse and Flinders areas of east Australia are from ref. 25, and three (grey-filled symbols) of the four ages from Ebor volcano are from ref. 26. All other ages for the Main Range, Tweed, Belmore, Ebor and Comboyne volcanoes are from this study. Errors (2σ) are smaller than the symbols or are illustrated by error bars. Robust, non-parametric regression of the time windows at 31–26 Myr, 26–23 Myr and 23–16 Myr yields the northward component of Australian plate motion, relative to a fixed hotspot reference frame, shown with 1σ uncertainties (see Methods for details).

- Mahoney, J.J. & Coffin, M.F.) *Am. Geophys. Union Geophys. Monogr.* **100**, 183–216 (1997).
9. Coffin, M.F. & Eldholm, O. Scratching the surface: estimating dimensions of large igneous provinces. *Geology* **21**, 515–518 (1993).
 10. Sharp, W.D. & Clague, D.A. 50-Ma initiation of Hawaiian–Emperor bend records major change in Pacific Plate motion. *Science* **313**, 1281–1284 (2006).
 11. Petterson, M.G. *et al.* Structure and deformation of north and central Malaita, Solomon Islands: Tectonic implications for the Ontong Java Plateau–Solomon arc collision, and for the fate of oceanic plateaus. *Tectonophysics* **283**, 1–33 (1997).
 12. Cooper, P.A. & Taylor, B. Polarity reversal in the Solomon Islands arc. *Nature* **314**, 428–430 (1985).
 13. Mann, P. & Taira, A. Global tectonic significance of the Solomon Islands and Ontong Java Plateau convergent zone. *Tectonophysics* **389**, 137–190 (2004).
 14. Morgan, W.J. Convection plumes in the lower mantle. *Nature* **230**, 42–43 (1971).
 15. Clouard, V. & Bonneville, A. How many Pacific hotspots are fed by deep-mantle plumes? *Geology* **29**, 695–698 (2001).
 16. Tarduno, J.A. *et al.* The Emperor seamounts: Southward motion of the Hawaiian hotspot plume in Earth's mantle. *Science* **301**, 1064–1069 (2003).
 17. Foulger, G.R. & Natland, J.H. Is 'hotspot' volcanism a consequence of plate tectonics? *Science* **300**, 921–922 (2003).
 18. DePaolo, D.J. & Manga, M. Deep origin of hotspots: The mantle plume model. *Science* **300**, 920–921 (2003).
 19. Courtillot, V., Davaille, A., Besse, J. & Stock, J. Three distinct types of hotspots in the Earth's mantle. *Earth Planet. Sci. Lett.* **205**, 295–308 (2003).
 20. Montelli, R. *et al.* Finite-frequency tomography reveals a variety of plumes in the mantle. *Science* **303**, 338–343 (2004).
 21. O'Neill, C., Müller, R.D. & Steinberger, B. On the uncertainties in hot spot reconstructions and the significance of moving hot spot reference frames. *Geochem. Geophys. Geosyst.* **6**, 10.1029/2004GC000784 (2005).
 22. McDougall, I. & Duncan, R.A. Age progressive volcanism in the Tasmanian seamounts. *Earth Planet. Sci. Lett.* **89**, 207–220 (1988).
 23. Duncan, R.A. & McDougall, I. in *Intraplate Volcanism in Eastern Australia and New Zealand* (ed. Johnson, R.W.) 43–53 (Cambridge Univ. Press, 1989).
 24. Koppers, A.A.P., Duncan, R.A. & Steinberger, B. Implications of a nonlinear $^{40}\text{Ar}/^{39}\text{Ar}$ age progression along the Louisville seamount trail for models of fixed and moving hot spots. *Geochem. Geophys. Geosyst.* **5**, doi:10.1029/2003GC000671 (2004).
 25. Cohen, B.E., Vasconcelos, P.M. & Knesel, K.M. $^{40}\text{Ar}/^{39}\text{Ar}$ constraints on the timing of Oligocene intraplate volcanism in southeast Queensland. *Aust. J. Earth Sci.* **54**, 105–125 (2007).
 26. Ashley, P.M., Duncan, R.A. & Feebrey, C.A. Ebor Volcano and Crescent Complex, northeastern New South Wales: Age and geological development. *Aust. J. Earth Sci.* **42**, 471–480 (1995).
 27. Wellman, P. Palaeomagnetism of two Mid-Tertiary basaltic volcanoes in Queensland Australia. *Proc. R. Soc. Qld* **86**, 147–153 (1975).
 28. McElhinny, M.W., Embleton, B.J.J. & Wellman, P. A synthesis of Australian Cenozoic palaeomagnetic results. *Geophys. J. R. Astr. Soc.* **36**, 141–151 (1974).
 29. Müller, R.D., Royer, J.-Y. & Lawver, L.A. Revised plate motions relative to the hotspots from combined Atlantic and Indian Ocean hotspot tracks. *Geology* **21**, 275–278 (1993).
 30. Ewart, A., Chappell, B.W. & Le Maitre, R.W. Aspects of the mineralogy and chemistry of the intermediate-silicic Cainozoic volcanic rocks of eastern Australia. Part 1: Introduction and geochemistry. *Aust. J. Earth Sci.* **32**, 359–382 (1985).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank M. Cloos, A. Ewart, M. Gasparon, A. Koppers, G. Rosenbaum and W. Sharp for comments; the Queensland and New South Wales Parks and Wildlife Services and various landowners for permission to undertake fieldwork on their properties; and A. Ewart and F.L. Sutherland for providing samples. Construction of the University of Queensland Argon Geochronology in Earth Sciences laboratory (UQ-AGES) was partially funded by the ARC; this project was funded through UQ-AGES contract research and an Australian Postgraduate award to B.E.C.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to K.M.K. (k.knesel@uq.edu.au) and B.E.C. (b.cohen@uq.edu.au).

METHODS

^{40}Ar – ^{39}Ar analytical methods. After removal of visibly weathered material, rock samples were crushed and sieved to ~ 0.3 – 2 mm, washed in distilled water in an ultrasonic bath for more than 30 min, and then washed in absolute ethanol. Grains were selected by hand picking under a binocular microscope; wherever possible, we selected fresh, inclusion-free, K-bearing phenocrysts (sanidine, anorthoclase, biotite, hornblende), 0.5 to 2 mm in diameter. Cloudy, friable and inclusion-rich crystals were avoided. We picked crystals with minor iron staining or minor opaque inclusions for samples where these features were unavoidable, particularly for anorthoclase crystals from metaluminous trachytes. Where phenocrysts were unavailable (typically for fine-grained mafic lavas, but also for two trachytes and one comendite), chips of visibly fresh, well-crystallized groundmass were analysed. Thin-section inspection revealed that all potassium-bearing phases are visibly unaltered, but some mafic groundmass samples contain variable amounts of altered olivine and/or greenish-brownish interstitial material indicative of minor deuteric alteration and/or surficial weathering.

The selected mineral grains and groundmass chips were loaded into Al disks, along with the neutron fluence monitor Fish Canyon Sanidine (prepared at the Berkeley Geochronology Center; age 28.02 ± 0.09 Myr (ref. 31)), following the geometry illustrated in ref. 32, and irradiated for 14 h in the CLICIT facility, a TRIGA-type reactor at Oregon State University. Irradiation correction factors are: $(^{36}\text{Ar}/^{37}\text{Ar})_{\text{Ca}} = (2.64 \pm 0.02) \times 10^{-4}$, $(^{39}\text{Ar}/^{37}\text{Ar})_{\text{Ca}} = (7.04 \pm 0.06) \times 10^{-4}$, and $(^{40}\text{Ar}/^{39}\text{Ar})_{\text{K}} = (8 \pm 3) \times 10^{-4}$. All ages are reported using the decay constants of ref. 33.

Mass spectrometric analyses were performed at the UQ-AGES laboratory. Before analysis, samples were baked out at $\sim 200^\circ\text{C}$ under vacuum for ~ 24 h. Most samples were analysed in duplicate by the incremental-heating method, using a continuous-wave Ar-ion laser with a 2-mm-wide defocused beam, following the procedures outlined in ref. 32 and 34. Usually single crystals or chips of groundmass were analysed, but in some cases clusters of two to three crystals (in one case eight) were combined in one aliquot. Optically clear sanidine crystals (including the Fish Canyon standards) do not couple well with the 480–540-nm Ar-ion laser used at UQ-AGES. These crystals were analysed by the total-fusion method, with one or two <0.2 mm grains of Zero Age Glass (ZAG) added to assist fusion, as outlined in ref. 32. We analysed between 8 and 10 aliquots of single sanidine crystals (rarely, clusters of two to three crystals) for each sample. Active gases released during laser heating were removed by a cryo-cooled trap (-138°C) and two C-50 SAES Zr-V-Fe getter pumps in the extraction line, and a third C-50 SAES Zr-V-Fe getter pump in the mass spectrometer. The clean gas sample was analysed for argon isotopes in a MAP-215-50 mass spectrometer. Air pipettes and full-system blanks were analysed before and after each sample. We corrected the data for mass discrimination, atmospheric contamination and nucleogenic interferences following the procedures in ref. 32, using MassSpec Version 7.527 for Macintosh Operating System 10 developed by A. Deino (Berkeley Geochronology Centre).

For each irradiation disk, we analysed 15 individual samples (either single crystals or two to three crystals combined) of Fish Canyon sanidines by the laser total-fusion method, yielding J factors listed in Supplementary Table 3. We calculated the mass spectrometer gain based on the analysis of an air pipette (1.72×10^{-13} moles ^{40}Ar) on the Faraday detector (~ 4 mV) equipped with a 10^{11} -ohm resistor, yielding a sensitivity of $\sim 4 \times 10^{-14}$ mol mV^{-1} . The same signal measured on a Balzers 217 Electron Multiplier, operated with a gain of $\sim 85,000$, produced a signal of ~ 3.5 nA ^{40}Ar , yielding a sensitivity of the order of 5×10^{-19} mol mV^{-1} .

^{40}Ar – ^{39}Ar systematics. In Supplementary Figs 1–5, we present ^{40}Ar – ^{39}Ar incremental-heating spectra, total-fusion results and isochrons for lavas and hypabyssal rocks from the Main Range, Tweed, Belmore, Ebor and Comboyne central volcanoes. Incremental-heating experiments generally yield near-ideal flat

spectra over more than 50% of the ^{39}Ar released. Isochron analysis demonstrates initial $^{40}\text{Ar}/^{36}\text{Ar}$ ratios indistinguishable at 95% confidence from the present-day value of 295.5 (ref. 35). These features are indicative of closed-system behaviour of argon, with no sign of argon losses, excess argon or recoil (ref. 36). Five samples from this study yield descending or slightly saddle-shaped incremental-heating spectra. Cast on isochron diagrams, these samples yield a single linear array indicative of a single age population (mean square weighted deviation <2.0), but with initial $^{40}\text{Ar}/^{36}\text{Ar}$ ratios above 295.5 at 95% confidence, consistent with excess argon trapped in incompletely degassed lavas and hypabyssal intrusions (ref. 37). This excess argon component, if undetected, would cause up to a 10% overestimate in age, even for mid-Cenozoic subaerial lavas. Only in two samples did incremental-heating produce slightly ascending spectra indicative of minor Ar losses due to weathering.

Published ^{40}Ar – ^{39}Ar ages and age–latitude regressions. In this work, we also use incremental-heating and total-fusion ^{40}Ar – ^{39}Ar results from 28 samples from our pilot study of the Fraser, Noosa, Glasshouse and Flinders volcanic areas in southeast Queensland²⁵ and seven incremental-heating analyses from the Ebor volcano in New South Wales²⁶. Reference 25 used the same analytical facility and laboratory procedures as this study, and the ages can be compared directly with the results reported here. Ages from ref. 26, however, were reported using a Fish Canyon Tuff biotite age of 27.7 Myr, with 1σ errors. To aid comparison with our ^{40}Ar – ^{39}Ar results, we have recalculated the ages of ref. 26 relative to a Fish Canyon Tuff sanidine age of 28.02 Myr (ref. 31), and have recalculated the errors to 2σ . The recalculated results are about 0.3 Myr older. Finally, the northward component of Australian plate motion, relative to a fixed hotspot reference, was derived by a non-parametric regression of the age versus latitude data for the time windows 31–26 Myr, 26–23 Myr and 23–16 Myr using the robust regression option in Isoplot version 3.00 (ref. 38). The robust regression makes no assumptions about the causes of scatter of the data from a straight line and requires no arbitrary decisions about what data should be included in the regression³⁸.

Seamount migration rates and seafloor bathymetry. The southward migration of volcanism for the Tasmanid and Lord Howe seamount tracks, shown at 5-Myr intervals in Fig. 2, was estimated using migration rates from this study for the time period 35–16 Myr ago and velocities from ref. 39 for the time range from 16 Myr to present. The oldest ^{40}Ar – ^{39}Ar total-fusion ages for the Tasmanid Seamounts are shown in Fig. 2 because they likely represent the best estimates of the eruption ages for the partially altered dredged volcanic samples. Seafloor bathymetry used in Fig. 2 is from ref. 40.

1. Renne, P. R. *et al.* Intercalibration of standards, absolute ages and uncertainties in $^{40}\text{Ar}/^{39}\text{Ar}$ dating. *Chem. Geol.* **145**, 117–152 (1998).
2. Vasconcelos, P. M., Onoe, A. T., Kawashita, K., Soares, A. J. & Teixeira, W. $^{40}\text{Ar}/^{39}\text{Ar}$ geochronology at the Instituto de Geociências, USP: Instrumentation, analytical procedures, and calibration. *Ann. Brazil. Acad. Sci.* **74**, 297–342 (2002).
3. Steiger, R. H. & Jäger, E. Subcommission on geochronology: Convention on the use of decay constants in geo- and cosmochemistry. *Earth Planet. Sci. Lett.* **36**, 359–362 (1977).
4. Vasconcelos, P. M. K–Ar and $^{40}\text{Ar}/^{39}\text{Ar}$ geochronology of weathering processes. *Annu. Rev. Earth Planet. Sci.* **27**, 183–229 (1999).
5. Nier, A. O. A redetermination of the relative abundances of the isotopes of carbon, nitrogen, argon and potassium. *Phys. Rev.* **77**, 789–793 (1950).
6. McDougall, I. & Harrison, T. M. *Geochronology and Thermochronology by the $^{40}\text{Ar}/^{39}\text{Ar}$ Method*. (Oxford Univ. Press, New York, 1999).
7. Kelley, S. Excess argon in K–Ar and Ar–Ar geochronology. *Chem. Geol.* **188**, 1–22 (2002).
8. Ludwig, K. R. *Isoplot Version 3.00: A Geochronological Toolkit for Microsoft Excel*. Berkeley Geochronology Center Spec. Publ., **4**, (Berkeley Geochronology Center, 2003).
9. Tregoning, P. Plate kinematics in the western Pacific derived from geodetic observations. *J. Geophys. Res.* **107**, ECV 7-1–8 (2002).
10. Petkovic, P. & Buchanan, C. *Australian Bathymetry and Topography Grid [Digital Dataset]* (Geoscience Australia, Canberra, 2002).

LETTERS

Thermochemical flows couple the Earth's inner core growth to mantle heterogeneity

Julien Aubert¹, Hagay Amit², Gauthier Hulot² & Peter Olson³

Seismic waves sampling the top 100 km of the Earth's inner core reveal that the eastern hemisphere (40° E–180° E) is seismically faster^{1,2}, more isotropic^{2,3} and more attenuating⁴ than the western hemisphere. The origin of this hemispherical dichotomy is a challenging problem for our understanding of the Earth as a system of dynamically coupled layers. Previously, laboratory experiments have established that thermal control from the lower mantle can drastically affect fluid flow in the outer core⁵, which in turn can induce textural heterogeneity on the inner core solidification front⁶. The resulting texture should be consistent with other expected manifestations of thermal mantle control on the geodynamo, specifically magnetic flux concentrations^{7,8} in the time-average palaeomagnetic field^{9,10} over the past 5 Myr, and preferred eddy locations¹¹ in flows imaged below the core–mantle boundary by the analysis of historical geomagnetic secular variation¹². Here we show that a single model of thermochemical convection and dynamo action can account for all these effects by producing a large-scale, long-term outer core flow that couples the heterogeneity of the inner core with that of the lower mantle. The main feature of this thermochemical ‘wind’ is a cyclonic circulation below Asia, which concentrates magnetic field on the core–mantle boundary at the observed location and locally agrees with core flow images. This wind also causes anomalously high rates of light element release in the eastern hemisphere of the inner core boundary, suggesting that lateral seismic anomalies at the top of the inner core result from mantle-induced variations in its freezing rate.

In the lower mantle, the double crossing of the post-perovskite phase transition detected by core-reflected seismic shear waves¹³ directly constrains local temperature gradients. Heat flow from the core to the mantle is found¹⁴ to be larger than average below Central America, where descending mantle currents induce regional heat flow anomalies of $\delta q = 40 \text{ mW m}^{-2}$ or larger, and lower than average below the central Pacific. These results support a thermal interpretation of the largest scales present in seismic shear velocity maps¹⁵, according to which fast regions correspond to cold deep subducted lithosphere^{14,16} (as beneath Central America), and extract more heat from the core, while slow regions correspond to warm thermochemical piles^{17,18} (as beneath the central Pacific), and extract less heat from the core. Cooling causes the inner core to grow at nominal rates of $0.3\text{--}0.9 \text{ mm yr}^{-1}$ (ref. 19), corresponding to 100 km of solidification within the past 100–300 Myr. As the large-scale lower mantle structure has changed little during that time¹⁸, a connection between its present-day pattern and the upper inner core heterogeneous properties is plausible.

The following numerical simulation shows how thermal mantle heterogeneity simultaneously affects the time-average outer core magnetic field structure, core flow, and the asymmetric structure

of the inner core. We use the same model of Boussinesq convection and dynamo action in a rotating spherical shell as in a previous study¹¹, except for the use of a co-density formulation²⁰ to describe both thermal and chemical buoyancy sources in the outer core in terms of a single co-density variable C (see Methods). We assume that the thermal and chemical perturbations have the same effective diffusivity κ because of turbulent mixing. At the core–mantle boundary, we assume zero chemical mass flux, so the mass anomaly flux $f = -\kappa \partial C / \partial r$ there (which combines thermal and chemical contributions) is related to the heat flow q through $f = \alpha q / C_p$, where r is the radial coordinate, and α and C_p are respectively the thermal expansivity and the specific heat of the liquid outer core. Thermal mantle control is modelled by imposing a heterogeneous heat flow pattern proportional to lowermost mantle seismic shear velocity anomalies¹⁵ (Supplementary Fig. 1), superimposed on a uniform background heat flow, as in earlier studies^{7,8}. The seismically inferred amplitude^{13,14} $\delta q = 40 \text{ mW m}^{-2}$ of the lateral heat flow variation corresponds to a mass anomaly flux variation $\delta f = 5 \times 10^{-10} \text{ kg m}^{-2} \text{ s}^{-1}$, using²¹ $\alpha = 10^{-5} \text{ K}^{-1}$ and $C_p = 800 \text{ J kg}^{-1} \text{ K}^{-1}$. This is of the same order of magnitude as the estimated homogeneous part of the mass anomaly flux in the core²¹ $f_0 = 2 \times 10^{-10} \text{ kg m}^{-2} \text{ s}^{-1}$, expressed per unit surface at the core–mantle boundary. Accordingly, we specify $\delta f / f_0 = 1$ in our model. At the inner core boundary, the co-density C is set to a constant value, and f is free to react to the convection. This thermodynamically consistent boundary condition (see Supplementary Information) allows for a spatially variable release of heat and light elements, and implies²⁰ lateral variations in the inner core growth rate, which is fastest where outer core downwellings bring relatively cold and chemically depleted liquid close to the inner core boundary. The simulation parameters are chosen so that the model produces a self-sustained magnetic field with dipole reversals. Because of the great disparity in the timescales of outer core and mantle dynamical processes, we consider only the time-average behaviour of the model. The flow is then dominated by a thermochemical wind balance¹¹ between the pressure gradient and the Coriolis and buoyancy forces.

The most prominent non-axisymmetric feature of the resulting thermochemical wind (Fig. 1) is a cyclonic (anticlockwise on a north polar view) flow column, parallel to the rotation axis, extending from the outer to the inner boundary, just outside the inner core tangent cylinder. The column is located below Asia, and is driven by mantle-induced lateral temperature gradients between the cold sub-Asian region and warmer African and Pacific regions. Two additional columns are located below the Pacific and Atlantic oceans. Strong helical polar vortices are generated inside the northern and southern parts of the tangent cylinder. The Asian cyclone is visible in the flow at the top of the free stream (Fig. 2d), and locally concentrates²² the outer boundary radial magnetic field (Fig. 2b). Similar signatures are found

¹Dynamique des Fluides Géologiques, ²Géomagnétisme, Institut de Physique du Globe de Paris, Université Paris-Diderot, INSU/CNRS, 4, Place Jussieu, 75252, Paris cedex 05, France.

³Department of Earth and Planetary Sciences, Johns Hopkins University, Baltimore, Maryland 21218, USA.

in the time-average core flow (Fig. 2c) and palaeomagnetic field (Fig. 2a), showing that these patterns could indeed result from the presence of such a cyclone in the Earth's outer core. Our model (Fig. 2b, d) further shows that the second persistent palaeomagnetic flux lobe below north America (Fig. 2a) is associated with a similar long-term cyclone, suggesting that the anticyclonic flow inferred in this region from historical geomagnetic secular variation (Fig. 2c) may be a transient. This interpretation is supported by changes in the instantaneous rotation direction of this vortex¹², occurring between 1840 and 1990.

The thermochemical wind flows efficiently extract heat and light elements from the equatorial belt of the inner boundary (Fig. 2e), while the polar vortex circulations in the tangent cylinder suppress the buoyancy extraction, yielding lower-than-average mass anomaly flux in the polar regions. Our model thus confirms the likelihood of a faster inner-core growth in the equatorial region, in line with earlier predictions^{6,23}. It also predicts an extra mass anomaly flux release and locally faster inner-core growth below southeastern Asia, where the sub-Asian cyclone brings colder, chemically depleted material from the outer boundary towards the inner boundary. This maximum coincides with the maximum perturbation in the isotropic seismic velocity of the upper inner core² (Fig. 2e), which is also representative of the large-scale pattern of anisotropic^{2,3} and first-order attenuation⁴ properties. This result is weakly sensitive to the time-average mass anomaly flux partition between the inner and outer boundaries, as shown (Fig. 2f) by a simulation in which the inner-boundary mass

anomaly flux is increased from 50% to 80% of the total, to better match current estimates (see Methods).

Solidification texturing appears to be the most likely micromechanism for explaining seismic heterogeneity below the inner core boundary, because the alternative process, deformational texturing, is too slow, typically requiring 1 Gyr timescales²³. Solidification experiments reveal the importance of interstitial solute flow rate in controlling fabric development in hexagonal-close-packed alloys²⁴. Slower freezing rates result in more widely spaced dendritic platelets with greater sensitivity to the solute flow direction, and produce a more textured solid through preferential *c*-axis orientation, whereas faster freezing rates inhibit this effect and result in a solid with more random platelet orientation. Seismic waves that sample the top of the inner core will therefore have anisotropic wave speed and attenuation in the more textured slow-growing regions, and will be more isotropic, faster on average and more attenuated²⁵ (due to scattering by an increased number of grain boundary crossings) in the less-textured fast-growing regions. If inner core textures are controlled by the processes just described, then the heterogeneous crystallization predicted by our dynamo model qualitatively agrees with the ensemble of seismic data²⁵. Our results therefore strongly suggest that the upper inner core has inherited its seismic heterogeneity through mantle-induced lateral variations in its growth rate.

Our interpretation requires a small inner core rotation with respect to the lower mantle over the past 100–300 Myr. Subject to a long term magnetic torque $\bar{\Gamma}_B$, viscous torque $\bar{\Gamma}_v$ and gravitational

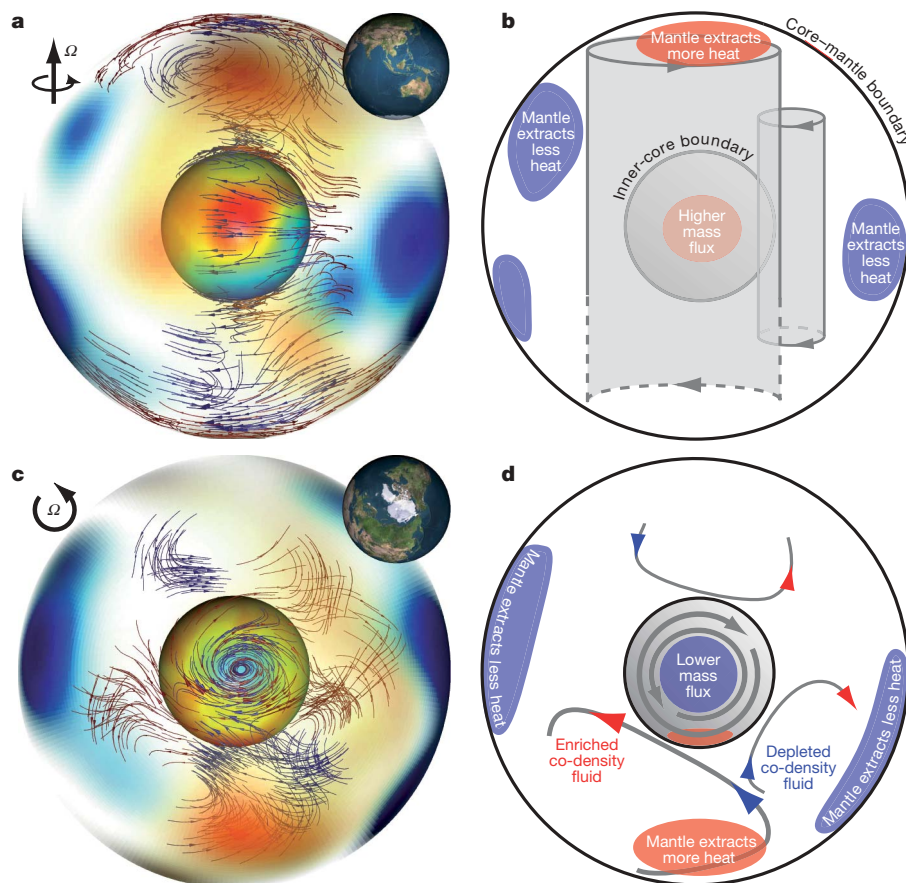


Figure 1 | Time-average flow structure of model case I. Parameters for this case are given in Methods. **a, c**, Equatorial and polar visualizations, respectively, of the time-average flow. Insets at upper right indicate the viewing angle. Insets at upper left show the rotation axis direction. The outer boundary is made selectively transparent and colour-coded according to the imposed outwards heat flow (a red patch, such as the dominant positive anomaly below Asia, means a larger-than-average heat flow; see also

Supplementary Fig. 1). The inner boundary is colour-coded according to the mass anomaly flux $f = -\kappa \partial C / \partial r$ extracted from the inner core (colour scheme as in Fig. 2e). Within the shell, flow streamlines are represented and colour-coded according to the local velocity along the cylindrical radius (blue streamlines represent downwells). **b, d**, Explanatory diagrams of the equatorial and polar views **a, c**. Red and blue arrowheads in **d** respectively represent flow up- and downwells.

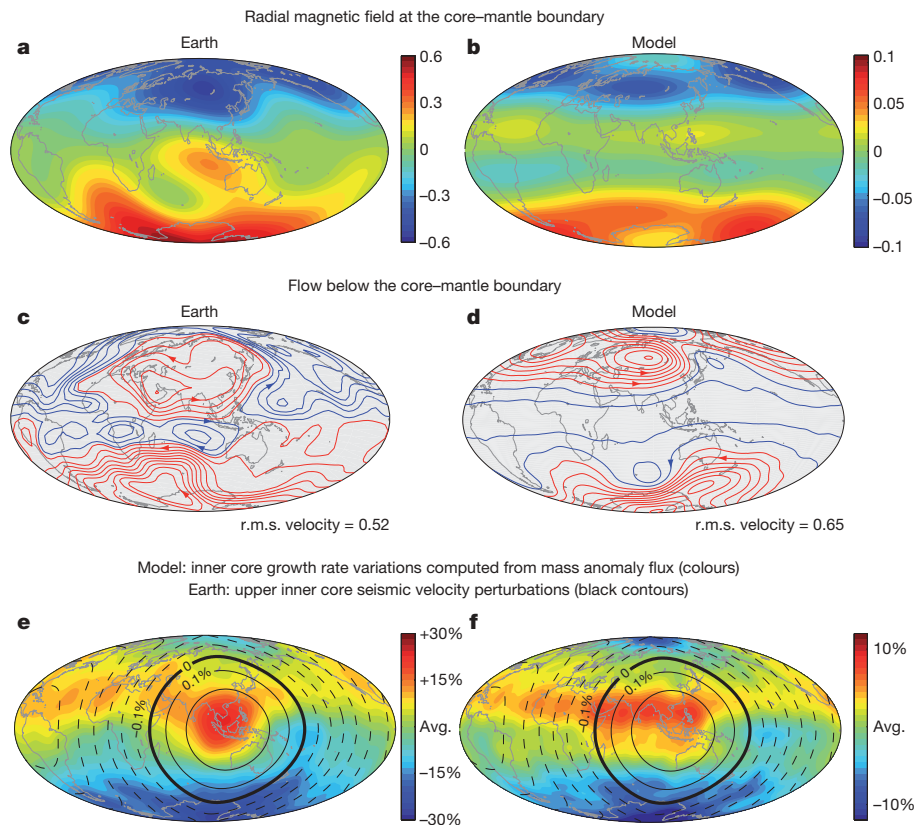


Figure 2 | Comparison between model and observations. **a**, Time-average palaeomagnetic field model¹⁰ at the core-mantle boundary over the normal polarity periods of the past 5 Myr. **b**, Time-average magnetic field at the outer boundary for case I, taken over the stable polarity periods of 0.7 Myr, and filtered to spherical harmonic degree and order 5. In **a** and **b**, the amplitude (colour scale) is normalized relative to the values predicted by a systematic scaling study²¹ (the predicted magnetic field strength in the Earth's core is 1 mT, see Supplementary Information). **c**, Streamlines of the time-average flow below the core-mantle boundary, obtained from geomagnetic secular variation inversions¹², for the period 1840–1990.

torque with coupling constant Γ_G , the inner core can be expected to reach a steady rotation rate²⁶ $\bar{\omega} = (\bar{\Gamma}_B + \bar{\Gamma}_v)/\tau\Gamma_G$, where τ is the relaxation time of the inner core towards its hydrostatic shape. The observation²⁷ of a 6-yr gravitational oscillation in length-of-day data implies that $\tau > 5$ yr and $\Gamma_G = 3 \times 10^{20}$ N m. Numerical dynamo simulations²⁶ with gravitational coupling further show that in the long term, the system tends to minimize $\bar{\Gamma}_B$, and that when the viscous torque is taken into account, both magnetic and gravitational torques independently counteract the viscous torque which entrains the inner core rotation. These results suggest that $\bar{\Gamma}_B + \bar{\Gamma}_v$ is only a fraction of the typical turbulent viscous torque, which we estimate by assuming a relative zonal flow velocity of the same order of magnitude (10^{-4} m s⁻¹) as at the core surface¹². Using equation (9) of ref. 28, this yields $\bar{\Gamma}_v = 2 \times 10^{12}$ N m, and a maximum inner core rotation of only 8° in 100 Myr, compatible with our results. On decadal time-scales, additional torsional oscillations are expected to disrupt the steady Taylor state of the outer core. For a 60-yr periodicity, magnetic torque fluctuations much larger than $\bar{\Gamma}_B$ can then cause the inner core to oscillate at a maximum angular velocity of²⁹ 0.1° yr⁻¹, within the range of present seismic inferences³⁰. Both processes would only smear, but not erase, the mantle signature on the inner core.

METHODS SUMMARY

We solve for fluid motion, thermochemical transport and magnetic induction in a spherical shell of aspect ratio $r_i/r_o = 0.35$, rotating with constant angular velocity Ω . The non-dimensional Navier–Stokes and induction equations are identical to those presented in an earlier study¹¹. The boundary conditions are rigid for the

d, Streamlines of the time-average flow at the top of the free stream for case I. In **c** and **d**, the root-mean-squared velocity is normalized relative to the values predicted by thermal wind scaling¹¹ (the predicted flow amplitude for the core is 15.8 km yr⁻¹, see also Supplementary Information). **e**, Colour map of relative anomalies of the mass anomaly flux extracted from the inner boundary for case I (or, equivalently, of inner-core growth rate anomalies). The black contours show an upper inner core seismic velocity perturbation model² expanded to spherical harmonic degree and order 1, the pattern of which is representative of anisotropic^{2,3} and attenuation⁴ seismic properties. **f**, Same as **e**, for case II (see Methods).

velocity field, and insulating for the magnetic field. The thermochemical buoyancy described by the co-density field C obeys the following evolution equation:

$$\frac{\partial C}{\partial t} + \mathbf{u} \cdot \nabla C = \frac{E}{\text{Pr}} \nabla^2 C + S_T - S_\chi \quad (1)$$

Here \mathbf{u} is the flow velocity field, and S_T and S_χ are respectively the volumetric thermal buoyancy source and chemical buoyancy sink terms. In model cases I and II, the relevant dimensionless parameters are set as follows: the Ekman number is $E = \nu/\Omega D^2 = 3 \times 10^{-4}$, where $D = r_o - r_i$ is the shell gap and ν is the fluid viscosity. The Prandtl number is $\text{Pr} = \nu/\kappa = 1$, and the magnetic Prandtl number is $\text{Pm} = \nu/\lambda = 2$, where λ is the magnetic diffusivity. The Rayleigh number based on the homogeneous mass anomaly flux²¹ f_o imposed at the outer boundary is $\text{Ra}_Q = g_o f_o / \rho \Omega^2 D^2$, where ρ is the fluid density and g_o is the gravity at the outer boundary. Model case I (Figs 1 and 2b, d, e) has $S_T + S_\chi = 0$ and $\text{Ra}_Q = 2 \times 10^{-4}$. Model case II (Fig. 2f) has $S_T + S_\chi = -2.1 \times 10^{-3}$ and $\text{Ra}_Q = 8.1 \times 10^{-5}$.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 9 November 2007; accepted 12 May 2008.

1. Niu, F. L. & Wen, L. X. Hemispherical variations in seismic velocity at the top of the Earth's inner core. *Nature* **410**, 1081–1084 (2001).
2. Tanaka, S. & Hamaguchi, H. Degree one heterogeneity and hemispherical variation of anisotropy in the inner core from PKP(BC)–PKP(DF) times. *J. Geophys. Res.* **102**, 2925–2938 (1997).
3. Yu, W. C. & Wen, L. X. Inner core attenuation anisotropy. *Earth Planet. Sci. Lett.* **245**, 581–594 (2006).
4. Cao, A. & Romanowicz, B. Hemispherical transition of seismic attenuation at the top of the Earth's inner core. *Earth Planet. Sci. Lett.* **228**, 243–253 (2004).

5. Sumita, I. & Olson, P. A laboratory model for convection in Earth's core driven by a thermally heterogeneous mantle. *Science* **286**, 1547–1549 (1999).
6. Bergman, M. I., Macleod-Silberstein, M., Haskel, M., Chandler, B. & Akpan, N. A laboratory model for solidification of Earth's core. *Phys. Earth Planet. Inter.* **153**, 150–164 (2005).
7. Olson, P. & Christensen, U. The time averaged magnetic field in numerical dynamos with non-uniform boundary heat flow. *Geophys. J. Int.* **151**, 809–823 (2002).
8. Gubbins, D., Willis, A. P. & Sreenivasan, B. Correlation of Earth's magnetic field with lower mantle thermal and seismic structure. *Phys. Earth Planet. Inter.* **162**, 256–260 (2007).
9. Johnson, C. L. & Constable, C. G. The time averaged geomagnetic field as recorded by lava flows over the past 5 Myr. *Geophys. J. Int.* **122**, 489–519 (1995).
10. Kelly, P. & Gubbins, D. The geomagnetic field over the past 5 million years. *Geophys. J. Int.* **128**, 315–330 (1997).
11. Aubert, J., Amit, H. & Hulot, G. Detecting thermal boundary control in surface flows from numerical dynamos. *Phys. Earth Planet. Inter.* **160**, 143–156 (2007).
12. Amit, H. & Olson, P. Time average and time dependent parts of core flow. *Phys. Earth Planet. Inter.* **155**, 120–139 (2006).
13. Lay, T., Hernlund, J., Garnero, E. J. & Thorne, M. S. A post-perovskite lens and D'' heat flux beneath the central Pacific. *Science* **314**, 1272–1276 (2006).
14. van der Hilst, R. et al. Seismostratigraphy and thermal structure of Earth's core-mantle boundary region. *Science* **315**, 1813–1817 (2007).
15. Masters, G., Laske, G., Bolton, H. & Dziewonski, A. in *Earth's Deep Interior: Mineral Physics and Tomography from the Atomic to the Global Scale* (eds Karato, S., Forte, A., Liebermann, R. C., Masters, G. & Stixrude, L.) 63–87 (AGU Monogr. Vol. 117, American Geophysical Union, Washington DC, 2000).
16. van der Hilst, R., Widiyantoro, S. & Engdahl, E. R. Evidence for deep mantle circulation from global tomography. *Nature* **386**, 578–584 (1997).
17. McNamara, A. K. & Zhong, S. J. Thermochemical structures beneath Africa and the Pacific Ocean. *Nature* **437**, 1136–1139 (2005).
18. Torsvik, T. H., Smethurst, M. A., Burke, K. & Steinberger, B. Large igneous provinces generated from the margins of the large low-velocity provinces in the deep mantle. *Geophys. J. Int.* **167**, 1447–1460 (2006).
19. Labrosse, S., Poirier, J. P. & Le Mouél, J. L. The age of the inner core. *Earth Planet. Sci. Lett.* **190**, 111–123 (2001).
20. Braginsky, S. I. & Roberts, P. H. Equations governing convection in Earth's core and the geodynamo. *Geophys. Astrophys. Fluid Dyn.* **79**, 1–97 (1995).
21. Christensen, U. & Aubert, J. Scaling properties of convection-driven dynamos in rotating spherical shells and application to planetary magnetic fields. *Geophys. J. Int.* **117**, 97–114 (2006).
22. Olson, P., Christensen, U. & Glatzmaier, G. A. Numerical modelling of the geodynamo: Mechanisms of field generation and equilibration. *J. Geophys. Res.* **104**, 10383–10404 (1999).
23. Yoshida, S., Sumita, I. & Kumazawa, M. Growth model of the inner core coupled with the outer core dynamics and the resulting elastic anisotropy. *J. Geophys. Res.* **101**, 28085–28103 (1996).
24. Bergman, M. I., Agrawal, S., Carter, M. & Macleod-Silberstein, M. Transverse solidification textures in hexagonal close-packed alloys. *J. Cryst. Growth* **255**, 204–211 (2003).
25. Cormier, V. Texture of the uppermost inner core from forward and back scattered seismic waves. *Earth Planet. Sci. Lett.* **258**, 442–453 (2007).
26. Buffett, B. A. & Glatzmaier, G. A. Gravitational braking of inner-core rotation in geodynamo simulations. *Geophys. Res. Lett.* **27**, 3125–3128 (2000).
27. Mound, J. E. & Buffett, B. A. Detection of a gravitational oscillation in length-of-day. *Earth Planet. Sci. Lett.* **243**, 383–389 (2006).
28. Aurnou, J. & Olson, P. Control of inner core rotation by electromagnetic, gravitational and mechanical torques. *Phys. Earth Planet. Inter.* **117**, 111–121 (2000).
29. Dumberry, M. Geodynamic constraints on the steady and time-dependent inner core axial rotation. *Geophys. J. Int.* **170**, 886–895 (2007).
30. Souriau, A. in *Treatise on Geophysics* Vol. 1, *Seismology and Structure of the Earth* (eds Dziewonski, A. & Romanowicz, B.) 655–693 (Elsevier, 2007).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements J.A. was supported by the SEDIT programme of CNRS-INSU. H.A. was supported by a Marie Curie intra-European grant. Numerical calculations were performed at the Service de Calcul Parallèle, IGP, and at IDRIS, France. We thank S. Tanaka for providing published data, and V. Cormier for discussions. This is IGP contribution 2369.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to J.A. (aubert@ipgp.jussieu.fr).

METHODS

We consider an electrically conducting, incompressible fluid in a spherical shell between radii r_i and r_o . The aspect ratio is $r_i/r_o = 0.35$, as in the Earth's core. The shell is rotating about an axis \mathbf{e}_z with constant angular velocity Ω . We model thermal and chemical convection in the Boussinesq approximation, and define the deviation temperature field T' and light element molar fraction field χ' with respect to the adiabatic temperature and well-mixed molar fraction. Both buoyancy effects are grouped into a co-density (or density anomaly) field²⁰ C such that:

$$C = \alpha \rho T' + \Delta \rho \chi' \quad (1)$$

Here α is the thermal expansion coefficient, ρ is the fluid density, and $\Delta \rho$ is the density difference between light elements and pure iron. The temperature and molar fraction fields are assumed to have the same diffusivity κ , due to turbulent mixing in the outer core. This allows us to write a single transport equation for the co-density C , which is solved numerically in a non-dimensional form, together with the magnetic induction equation for the solenoidal magnetic field \mathbf{B} in the magnetohydrodynamic approximation, and the Navier–Stokes and thermochemical transport equations for the incompressible velocity field \mathbf{u} , and pressure P :

$$\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \nabla \mathbf{u} + 2\mathbf{e}_z \times \mathbf{u} + \nabla P = \text{Ra}_Q \frac{\mathbf{r}}{r_o} C + (\nabla \times \mathbf{B}) \times \mathbf{B} + E \nabla^2 \mathbf{u} \quad (2)$$

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times (\mathbf{u} \times \mathbf{B}) + \frac{E}{\text{Pm}} \nabla^2 \mathbf{B} \quad (3)$$

$$\frac{\partial C}{\partial t} + \mathbf{u} \cdot \nabla C = \frac{E}{\text{Pr}} \nabla^2 C + S_T - S_\chi \quad (4)$$

$$\nabla \cdot \mathbf{u} = 0 \quad (5)$$

$$\nabla \cdot \mathbf{B} = 0 \quad (6)$$

Time is scaled with the inverse of the rotation rate Ω^{-1} . Length is scaled with the shell gap $D = r_o - r_i$. The velocity is scaled with ΩD . The magnetic induction is scaled with $(\rho \mu)^{1/2} \Omega D$, where μ is the magnetic permeability of the fluid. The co-density is scaled with $|f_o|/\Omega D$, where f_o is the amplitude of the spherically symmetric part of the mass anomaly flux $\mathbf{f} = -\kappa \nabla C$ imposed at the outer boundary of the model.

At both boundaries, the boundary conditions for velocity are rigid, and insulating for the magnetic field. Although treating the inner core as an insulator is non-physical, the influence of inner-core conductivity is insignificant³¹, except for inner-core differential rotation, which is however not allowed in the present simulations.

At the outer boundary of the model, the mass anomaly flux is prescribed with an homogeneous part f_o and an heterogeneous part proportional to the lower mantle seismic shear wave tomography¹⁵ SB4L18 (see Supplementary Fig. 1), truncated at spherical harmonic degree and order 9, and scaled to have a peak-to-peak magnitude of $\delta f = f_o$. At the inner boundary, the co-density is fixed (see Supplementary Information for a justification of this condition).

The Ekman number is $E = \nu/\Omega D^2 = 3 \times 10^{-4}$, where ν is the fluid viscosity. The Prandtl number is $\text{Pr} = \nu/\kappa = 1$, and the magnetic Prandtl number is $\text{Pm} = \nu/\lambda = 2$, where λ is the magnetic diffusivity. The Rayleigh number based on mass anomaly flux²¹ is $\text{Ra}_Q = g_o f_o / \rho \Omega^3 D^2$, where g_o is the gravity at the outer boundary. In equation (4), S_T is a source term representing the cooling over time of the reference adiabatic thermal state, and S_χ is a sink term representing the enrichment over time of the reference chemical state. Adjusting S_T and S_χ influences the relative strength of the inner-boundary and outer-boundary originated mass anomaly fluxes on time average. Simulations are carried out for two cases with approximately equal convective power: case I (Figs 1 and 2b, d, e), where $S_T + S_\chi = 0$ and $\text{Ra}_Q = 2 \times 10^{-4}$, for which scaling laws already exist in the literature^{11,21}, and a geophysically more relevant³² case II (Fig. 2f) where $S_T + S_\chi = -2.1 \times 10^{-3}$ and $\text{Ra}_Q = 8.1 \times 10^{-5}$. In case II, the inner-boundary originated mass anomaly flux represents 80% of the total.

The numerical implementation MAGIC³¹ is used in this study. A Chebyshev polynomials expansion is used in the radial direction with 65 radial grid points. A spherical harmonics expansion is used in the lateral directions, up to maximum degree and order 85.

The model is time-averaged during 4.4 magnetic diffusion times D^2/λ . Using $D = 2,200$ km and $\lambda \approx 1 \text{ m}^2 \text{ s}^{-1}$ for iron in the Earth's core³³, this is equivalent to a period of about 700 kyr. The time-average magnetic field is taken during the run periods of stable dipole.

31. Wicht, J. Inner-core conductivity in numerical dynamo simulations. *Phys. Earth Planet. Inter.* **132**, 281–302 (2002).
32. Lister, J. R. & Buffett, B. A. The strength and efficiency of thermal and compositional convection in the geodynamo. *Phys. Earth Planet. Inter.* **91**, 17–30 (1995).
33. Secco, R. A. & Shloessin, H. H. The electrical resistivity of solid and liquid Fe at pressures up to 7 GPa. *J. Geophys. Res.* **94**, 5887–5894 (1989).

LETTERS

Escape from adaptive conflict after duplication in an anthocyanin pathway gene

David L. Des Marais¹ & Mark D. Rausher¹

Gene duplications have been recognized as an important source of evolutionary innovation and adaptation since at least Haldane¹, and their varying fates may partly explain the vast disparity in observed genome sizes². The expected fates of most gene duplications involve primarily non-adaptive substitutions leading to either non-functionalization of one duplicate copy or subfunctionalization³, neither of which yields novel function. A significant evolutionary problem is thus elucidating the mechanisms of adaptive evolutionary change leading to evolutionary novelty. Currently, the most widely recognized adaptive process involving gene duplication is neo-functionalization (NEO-F), in which one copy undergoes directional selection to perform a novel function after duplication⁴. An alternative, but understudied, adaptive fate that has been proposed is escape from adaptive conflict (EAC), in which a single-copy gene is selected to perform a novel function while maintaining its ancestral function^{5,6}. This gene is constrained from improving either novel or ancestral function because of detrimental pleiotropic effects on the other function. After duplication, one copy is free to improve novel function, whereas the other is selected to improve ancestral function. Here we first present two criteria that can be used to distinguish NEO-F from EAC. Using both tests for positive selection and assays of enzyme function, we then demonstrate that adaptive evolutionary change in a duplicated gene of the anthocyanin biosynthetic pathway in morning glories (*Ipomoea*) is best interpreted as EAC. Finally, we argue that this phenomenon likely occurs more often than has been previously believed and may thus represent an important mechanism in generating evolutionary novelty.

Novel gene function associated with gene duplication can arise in either of two ways. According to the NEO-F model, novel function arises after gene duplication, with one copy maintaining ancestral function, whereas the second copy is selected to perform a new function. By contrast, under EAC, novel function arises first in the single-copy ancestral gene, which results in a reduction in that copy's ability to perform the original function. After duplication, each copy is free to specialize on either the original or the novel function and improve those functions. Although both of these processes have been recognized for at least two decades as alternative possibilities^{4,6}, we are aware of only one previous attempt to distinguish between them⁷.

Two criteria may be used to distinguish between NEO-F and EAC. First, under EAC adaptive change occurs in both duplicate copies, whereas under NEO-F only one copy undergoes adaptive change because purifying selection acts to maintain ancestral function in one copy. Second, under EAC ancestral function is improved, whereas under NEO-F it is not. Here we apply both of these criteria to distinguish between NEO-F and EAC as explanations for adaptive change in duplicated copies of the anthocyanin biosynthetic pathway gene dihydroflavonol-4-reductase (DFR).

Plant dihydroflavonol-4-reductases (EC number 1.1.1.219) function most conspicuously in the reduction of several flavonoid

precursors of anthocyanin pigments and their related phytoalexins, though the full scope of DFR function is unknown. The widely distributed 3-hydroxyanthocyanidins pelargonidin, cyanidin and delphinidin are primarily responsible for red, purple and blue flowers in angiosperms and are the downstream products of DFR activity on dihydrokaempferol (DHK), dihydroquercetin (DHQ) and dihydromyricetin (DHM), respectively⁸. DFR from several taxa is able to reduce the flavonols naringenin and eriodictiol to produce the flavan-4-ols, which are themselves precursors of phlobaphenes and of two rare anthocyanidins, apigeninidin and luteolinidin^{9,10}. In the common morning glory, *Ipomoea purpurea*, DFR is present as a small, tandemly arrayed three-gene family spanning approximately 17 kilobases (ref. 11). All three copies are expressed in *I. purpurea* (although *DFR-A* and *DFR-C* are expressed in overlapping but fewer tissues relative to *DFR-B*) and have the conserved intron/exon structure observed across eudicots.

Phylogenetic analysis of *DFR* in the *Ipomoea* gene family and closely related single-copy taxa shows that each *DFR* copy identified in *I. purpurea*, A, B and C, forms a clade (Fig. 1). Each clade of orthologues, the entire *DFR* gene family, *DFR* from all Convolvulaceae and *DFR* from all Solanales constitute well-supported clades. Sister to the three-copy species clade is a large clade of morning glories characterized by a single copy of *DFR* and represented here by *Evolvulus glomeratus*; most Solanaceae species are also single-copy and our results suggest that there was a lineage specific duplication in *Petunia*. The gene tree also reveals two separate *DFR* duplication events in the Convolvulaceae: the first gave rise to *DFR-B* and another lineage that experienced a second duplication event creating the *DFR-A* and *DFR-C* copies.

Codon-based models of sequence evolution¹² indicate that single-copy *DFRs* have historically experienced purifying selection (dN/dS (ω) = 0.13; Fig. 1), as did the base of the clade containing the Convolvulaceae *DFR* gene family (dN/dS = 0.104). After the first gene duplication, the lineage subtending *DFR-B* has a ratio of replacement to synonymous substitutions statistically indistinguishable from the single copy *DFRs* (dN/dS = 0.183). By contrast, this ratio has increased on the lineage subtending the *DFR-A/C* clade (dN/dS = infinite). A likelihood model infers that 18 replacement and no synonymous substitutions occurred at the base of the A/C clade. The probability of this pattern occurring under neutrality is P = 0.004 (see Supplementary Methods), indicating the action of repeated positive selection along this branch after the first duplication. Non-synonymous substitution was also significantly elevated above background on the branch subtending the *DFR-A* clade (dN/dS = 0.842, P = 0.009) and marginally so on the branch subtending the *DFR-C* clade (dN/dS = 0.29, P = 0.06), although a branch-sites test fails to detect positive selection on either of these (P = 0.21 and P > 0.9, respectively). Changes immediately after the second duplication thus hint at positive selection, but do not definitively reveal it.

¹Department of Biology and University Program in Genetics and Genomics, Box 90338, Duke University, Durham, North Carolina 27708-0338, USA.

Subsequent non-synonymous substitution on copies in the B and the A/C clades varies slightly from background rates but still bears the signature of purifying selection. A clade-sites test for the *DFR-A/C* clade as a whole compared with *DFR-B* as a whole reveals that most codon sites in both clades remain under strong purifying selection (Fig. 1). It thus appears that there was a burst of adaptation in the A and C copies very soon after the first, and possibly after the second, duplication. After this burst, selection once again became primarily purifying.

We chose *Solanum lycopersicon* and *E. glomeratus* as representative 'pre-duplication' species and *I. purpurea* and *Convolvulus arvensis* as representative 'post-duplication' species to determine whether enzyme function changed after duplication. Protein products of each *DFR* copy from these four taxa were assayed for enzymatic activity on three common *DFR* substrates (dihydrokaempferol, dihydroquercetin and dihydromyricetin) and two substrates less commonly reduced by plant *DFR* enzymes (naringenin and eriodictiol).

Three major patterns arise from this analysis. First, the pre-duplication copies exhibit moderate activity on DHK, DHQ and DHM (Fig. 2c–e), very low activity on eriodictiol (Fig. 2b) and minimal activity on naringenin (Fig. 2a). Second, the post-duplication A and C copies exhibit essentially no activity on any of the five substrates (Fig. 2). Moreover, *DFR-A* and *-C* activities on eriodictiol, DHK, DHQ and DHM are significantly reduced compared with activity in the pre-duplication copies (see Supplementary Methods for statistical analysis). Third, activity on all five substrates is higher for the post-duplication *DFR-B* copies than for the pre-duplication copies, all but DHQ significantly so even after a Bonferroni correction for multiple comparisons (see Supplementary Methods). The increased *DFR-B* activities on the molecularly smaller naringenin (4.6-fold), eriodictiol

(6.1-fold) and DHK (7.9-fold) are substantially greater than the increases on DHQ (1.4-fold) and DHM (1.4-fold).

These results indicate that adaptive change in *DFR* immediately subsequent to the first duplication is best explained as escape from adaptive conflict. Adaptive change apparently occurred in both copies after the first duplication, consistent with expectations under EAC. Adaptive change in the A/C copy is indicated by the signature of repeated positive selection in the lineage subtending the A/C clade. Although such a signature of positive selection was not detected on the B copy, enzyme functional analysis indicates that adaptive substitution likely occurred. In particular, after duplication, enzyme activity on ancestral flavonoid substrates increased substantially. Because most mutations are believed to have detrimental effects on enzyme stability or activity¹³, we believe that neutral evolution in the B-lineage can be rejected in favour of positive selection that is undetected at the sequence level, possibly because it involved few substitutions.

Further evidence for adaptive change along the B-lineage is provided by an analysis of the substitutions that occurred at its base. The crystal structure of *DFR* from *Vitis vinifera* shows that Asn 133 lines

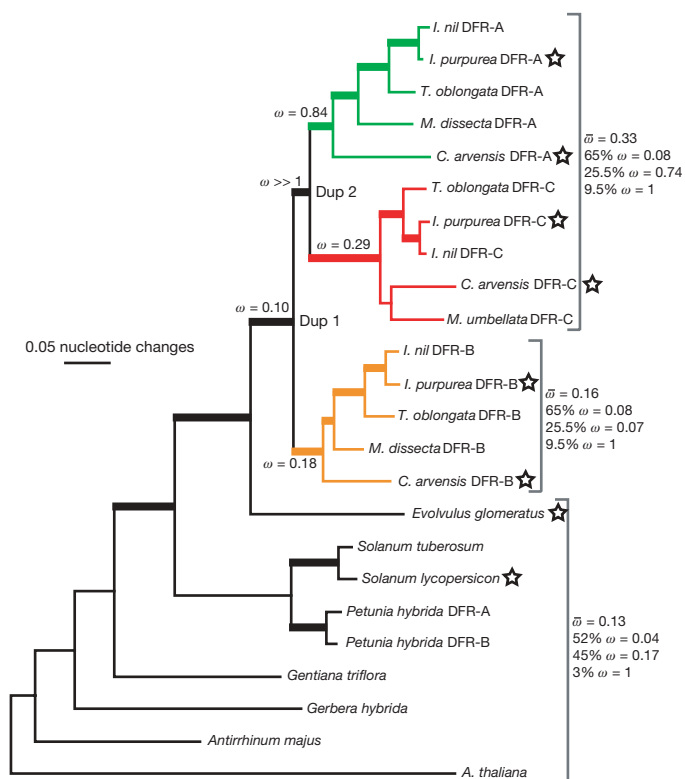


Figure 1 | Gene tree of sampled *DFR* copies from *Convolvulaceae* and outgroup taxa. Topology is identical under bayesian, maximum likelihood and maximum parsimony criteria. Thick branches indicate clades with strong statistical support (bayesian posterior probability greater than 0.95). Starred genes are those included in enzymatic studies. Results of selection tests are reported for tested branches, as are the proportion of sites under particular selective regimes for the *DFR-B*, *DFR-A/C* and outgroup clades. ω is the ratio of non-synonymous to synonymous nucleotide substitutions.

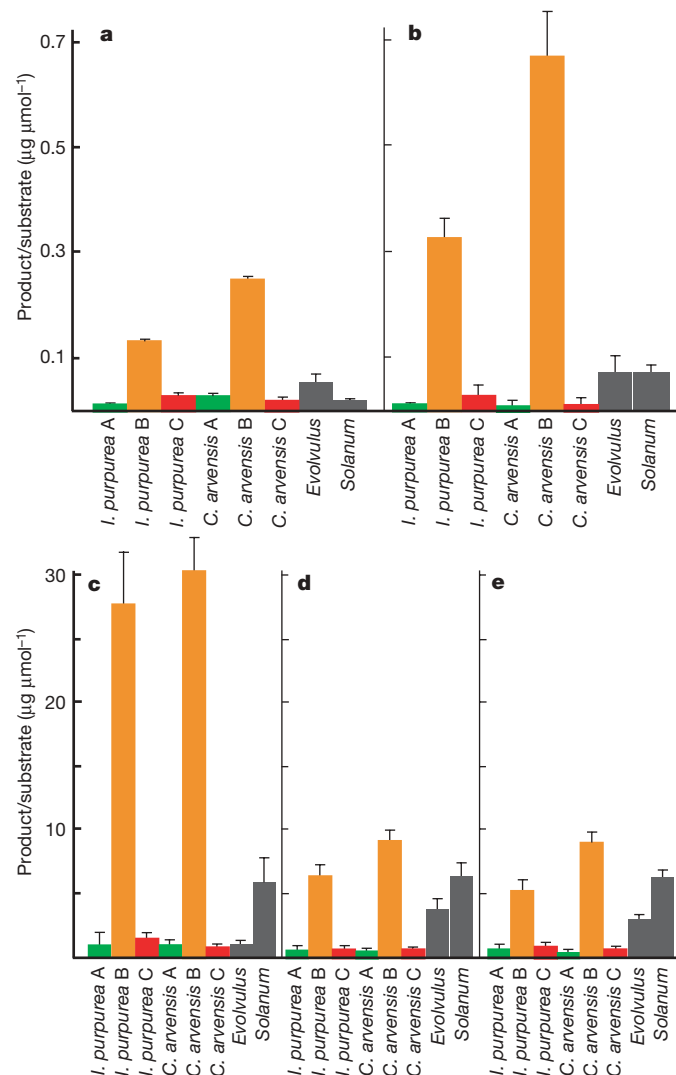


Figure 2 | Results of enzyme activity assays. Substrates are naringenin (a), eriodictiol (b), dihydrokaempferol (c), dihydroquercetin (d) and dihydromyricetin (e). These correspond to products apigeninidin and luteolinidin (3-deoxyanthocyanidins, top chart) and pelargonidin, cyanidin and delphinidin (bottom chart). Results are in micrograms of product per micromole of substrate. Colours reflect coloured orthologue clades as in Fig. 1. Bars are one standard error; $n = 4$ for each assay.

the substrate-binding pocket and forms a hydrogen bond with the 4' hydroxyl (and, if present, the 3' hydroxyl) of the substrate B-ring¹⁴. DFR from maize¹⁰ and *Pyrus*⁹ has activity on naringenin, eriodictiol, DHK and DHQ (DHM activity was not assayed in either of these studies). Both of these DFR variants have asparagine in the 133 position; in fact, nearly all known DFR genes from angiosperms have Asn133. Our phylogenetic reconstruction reveals that an Asn133 Asp substitution occurred before the divergence of the Gentianales and Solanales (Supplementary Fig. 1); this aspartate is retained in our two assayed single-copy species and correlates with relatively poorer flavonoid function. DFR-A copies from different species have different non-Asn amino acids at this site, and all DFR-C copies have the hydrophobic isoleucine; substituting the hydrophobic leucine at residue 133 has been shown to eliminate flavonoid activity in *Gerbera* DFR¹⁵. After gene duplication, the Asn133 variant re-evolved in the DFR-B lineage, coinciding with the improvement of DFR-B activity on the five substrates examined. Given this pattern, it is tempting to suggest that the original substitution of Asp for Asn at the base of the Gentianales and Solanales reflects the acquisition by the single-copy DFR of the unknown second function, and that the reverse transition in the DFR-B lineage reflects loss of this function in this lineage. Moreover, the failure to regain Asn in DFR-A and DFR-C is consistent with the loss of the ability of these copies to metabolize dihydroflavonols.

The improvement of ancestral function in the DFR-B lineage satisfies the second criterion for the operation of EAC. Apparently, the ability of DFR to reduce dihydroflavonols was constrained in the ancestral single-copy gene, and this constraint was released when the first duplication occurred.

Ideally, if EAC explains evolutionary change in duplicated copies of DFR, one should be able to demonstrate that the different copies perform different functions and that both of these functions were performed by the ancestral single-copy DFR, albeit at reduced efficiency. Unfortunately, we are unable to confirm this expectation because we have not yet identified what function(s) the DFR-A and -C copies perform. Although it is clear that these copies have lost the ability to perform one ancestral function of DFR (reduction of dihydroflavonols), we do not have direct evidence that they perform an ancestral rather than a novel function. Nevertheless, the improvement of flavonoid function in the DFR-B copies points to this function being constrained in the ancestral single-copy gene. Escape from adaptive conflict is thus clearly a feature of DFR evolution.

A potential alternative explanation for our results is that the activity spectrum of DFR-B reflects the ancestral state of DFR, and that the loss of ability to metabolize naringenin and eriodictiol, as well as reduced function on DHK, DHQ and DHM, in *E. glomeratus* and *S. lycopersicon* reflects independent losses of activity. However, an ancestral state reconstruction of substrate specificity does not support this interpretation: an analysis using all species for which activity on all five flavonoid substrates were tested indicates that although ability to use all five substrates is deeply ancestral in the angiosperms, the ability to use naringenin and eriodictiol was significantly reduced on the lineage leading to *Solanum*, *Evolvulus* and *Ipomoea* (Supplementary Fig. 2). Loss of function on these two substrates occurred before the DFR duplication in *Ipomoea*. This pattern is consistent with the EAC hypothesis: the reconstruction demonstrates first a reduction of ability to metabolize dihydroflavonols, naringenin and eriodictiol, corresponding to when the single-copy DFR gained an additional (currently unknown) function; and then an increase in the ability to metabolize these compounds after duplication, corresponding to release from adaptive constraint.

Our analysis illustrates the larger issue that many published examples of NEO-F may actually represent EAC. There are some cases of NEO-F that by the two criteria used in our analysis are better explained by NEO-F than by EAC^{16,17}. However, there are also many purported cases of NEO-F that have not characterized ancestral function before duplication, that have not determined whether adaptive

evolution occurred in one or in both duplicate copies subsequent to duplication, and that have not determined whether either copy exhibits improvement of known ancestral function (for example, the origin of vertebrate MHC¹⁸ and plant R-gene alleles¹⁹). In these cases, EAC cannot be excluded. Consequently, it remains unclear whether the absence of reported cases of EAC indicates a true rarity of its occurrence, or just that much of what is considered NEO-F is actually EAC.

Unfortunately, the literature is silent on the expected relative frequencies of these two processes. Although some models predict that the relative importance of NEO-F and subfunctionalization depends on factors such as population size^{20,21}, these results cannot necessarily be extrapolated to either a comparison of NEO-F and EAC or of EAC and subfunctionalization. Although EAC resembles subfunctionalization in that it results in the partitioning of ancestral functions, subfunctionalization is envisioned as involving only neutral or slightly deleterious mutations²², whereas EAC involves primarily adaptive substitutions. Similarly, although NEO-F involves adaptive mutations in one copy of a duplicate, EAC involves adaptive mutations in both copies. Neither of these features of EAC is incorporated into any existing models of the fate of duplicate genes. Intuitively, we suspect that the probability of preservation of duplicate copies is higher for EAC than for NEO-F, given that the duplicated gene has already acquired multiple functions, because mutations in either copy can start the EAC process. However, the overall occurrence of EAC compared with NEO-F will also depend upon the fraction of duplicated genes that have acquired new functions before duplication, a quantity for which no information exists.

That the DFR-B copy also evolved the capability of reducing naringenin and eriodictiol is surprising because the presumed downstream products of these reactions are rare in angiosperms and unreported from the Solanales²³. It is therefore doubtful that this new capacity is actually used in post-duplication *Ipomoea* species. Nevertheless, 3-deoxyanthocyanins and their polymers have been reported in some crops to have insecticidal and anti-fungal properties²⁴, which suggests that the ability of DFR-B to synthesize their precursors may constitute an exaptation for the evolution of such resistance. Under this interpretation, improvement of DFR-B activity on the more common substrates DHK, DHQ and DHM may have been the adaptive change whereas improvement in flavonol reduction was a byproduct. Interestingly, this result indicates that the evolution of novel function may sometimes be a byproduct of the evolution in enhanced ancestral activity.

METHODS SUMMARY

We used PCR to isolate DFR copies from species sampled across the Convolvulaceae and Solanaceae in an effort to identify when on the known species phylogeny the duplication events occurred. Gene trees were reconstructed using maximum parsimony, maximum likelihood and bayesian criteria, and branch support was assessed by bootstrapping and by examining posterior probabilities of nodes. To determine whether changes in selection pressures followed inferred duplication events, we used codon-based models of substitution as implemented in PAML. Specifically, we performed lineage and clade-based tests of each DFR copy's subtending lineage and entire clade, as well as branch-sites tests and clade-sites tests.

We subcloned each DFR copy identified in *S. lycopersicon*, *E. glomeratus*, *C. arvensis* and *I. purpurea* into bacterial overexpression vectors and then expressed these constructs in *Escherichia coli* BL-21 cell lines. After protein extraction, each enzyme was assayed for substrate specificity on five compounds: DHK, DHQ, DHM, naringenin and eriodictiol. The products of these reactions were purified and then quantified spectrophotometrically.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 15 November 2007; accepted 14 May 2008.

Published online 25 June 2008.

1. Haldane, J. B. S. The part played by recurrent mutation in evolution. *Am. Nat.* **67**, 5–19 (1933).

2. Hahn, M. W., De Bie, T., Stajich, J. E., Nguyen, C. & Cristiani, N. Estimating the tempo and mode of gene family evolution from comparative genomic data. *Genome Res.* **15**, 1153–1160 (2005).
3. Lynch, M. & Conery, J. S. The evolutionary fate and consequences of duplicate genes. *Science* **290**, 1151–1155 (2000).
4. Ohno, S. *Evolution by Gene Duplication* (Springer, Berlin, 1970).
5. Hughes, A. L. The evolution of functionally novel proteins after gene duplication. *Proc. R. Soc. Lond. B* **256**, 119–124 (1994).
6. Piatigorsky, J. & Wistow, G. The recruitment of crystallins: new functions precede gene duplication. *Science* **252**, 1078–1079 (1990).
7. Hittinger, C. T. & Carroll, S. B. Gene duplication and the adaptive evolution of a classic genetic switch. *Nature* **449**, 677–681 (2007).
8. Schwartz-Sommer, Z. *et al.* Influence of transposable elements on the structure and function of the *A1* gene of *Zea mays*. *EMBO J.* **6**, 287–294 (1987).
9. Fischer, T. C., Halbwirth, H., Meisel, B., Stich, K. & Forkman, G. Molecular cloning, substrate specificity of the functionally expressed dihydroflavonol 4-reductases from *Malus domestica* and *Pyrus communis* cultivars and the consequences for flavonoid metabolism. *Arch. Biochem. Biophys.* **412**, 223–230 (2003).
10. Halbwirth, H., Martens, S., Wienand, U., Forkmann, G. & Stich, K. Biochemical formation of anthocyanins in silk tissue of *Zea mays*. *Plant Sci.* **164**, 489–495 (2003).
11. Inagaki, Y. *et al.* Genomic organization of the genes encoding dihydroflavonol 4-reductase for flower pigmentation in the Japanese and common morning glories. *Gene* **226**, 181–188 (1999).
12. Yang, Z. PAML, a program for phylogenetic analysis by maximum likelihood. *Comput. Appl. Biosci.* **13**, 555–556 (1997).
13. DePristo, M. A., Weinreich, D. M. & Hartl, D. L. Missense meanderings in sequence space: a biophysical view of protein evolution. *Nature Rev. Genet.* **6**, 678–687 (2005).
14. Petit, P. *et al.* Crystal structure of grape dihydroflavonol 4-reductase, a key enzyme in flavonoid biosynthesis. *J. Mol. Biol.* **368**, 1345–1357 (2007).
15. Johnson, E. T. *et al.* Alteration of a single amino acid changes the substrate specificity of dihydroflavonol 4-reductase. *Plant J.* **25**, 325–333 (2001).
16. Benderoth, M. *et al.* Positive selection driving diversification in plant secondary metabolism. *Proc. Natl Acad. Sci. USA* **103**, 9118–9123 (2006).
17. Zhang, J., Zhang, Y.-P. & Rosenberg, H. F. Adaptive evolution of a duplicated pancreatic ribonuclease gene in a leaf-eating monkey. *Nature Genet.* **30**, 411–415 (2002).
18. Li, W.-H. *Molecular Evolution* (Sinauer, Sunderland, Massachusetts, 1997).
19. Michelsmore, R. W. & Meyers, B. C. Clusters of resistance in plants evolve by divergent selection and a birth-and-death process. *Genome Res.* **8**, 1113–1130 (1998).
20. Lynch, M. & Force, A. The probability of duplicate gene preservation by subfunctionalization. *Genetics* **154**, 459–473 (2000).
21. Lynch, M., O'Hely, M., Walsh, B. & Force, A. The probability of preservation of a newly arisen gene duplicate. *Genetics* **159**, 1789–1804 (2001).
22. Force, A. *et al.* Preservation of duplicate genes by complementary, degenerative mutations. *Genetics* **151**, 1531–1545 (1999).
23. Iwashina, T. The structure and distribution of the flavonoids in plants. *J. Plant Res.* **113**, 287–299 (2000).
24. Nicholson, R. L., Kollipara, S. S., Vincent, J. R., Lyons, P. C. & Cadena-Gomez, G. Phytoalexin synthesis by the sorghum mesocotyl in response to infection by pathogenic and nonpathogenic fungi. *Proc. Natl Acad. Sci. USA* **84**, 5520–5524 (1987).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank E. Grotewold and B. Winkel for technical advice, and R. Hopkins for implementation of statistical models. This work was supported by National Science Foundation funding. D.L.D. was supported in part by a National Institutes of Health training grant to Duke University's University Program in Genetics and Genomics.

Author Contributions D.L.D. and M.D.R. designed the research; D.L.D. performed the research and the analyses; D.L.D. and M.D.R. wrote the paper.

Author Information The new DNA sequences reported here are deposited in GenBank under accession numbers EU189072–EU189082. Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to D.L.D. (dld3@duke.edu) or M.D.R. (mrausher@duke.edu).

METHODS

Plant materials and extractions. We used the phylogeny of Stefanovic *et al.*²⁵ to target our sampling to a set of species that would help determine the timing of evolutionary transitions from a single-copy *DFR* ancestor to the three-copy family observed in *I. purpurea* and *I. nil*¹¹. We focused on five species representing the phylogenetic diversity of the Convolvulaceae: *I. purpurea*, *C. arvensis*, *E. glomeratus*, *Merremia dissecta* and *Merremia umbellata*, and *S. lycopersicon* from the sister family Solanaceae. All plants were grown from seed. DNA and RNA were extracted from freshly harvested buds, leaves and stem tissue using the DNeasy Plant Mini kit and RNeasy kit (Qiagen), respectively, according to the manufacturer's protocol.

Cloning and sequencing of *DFR* copies. To isolate *DFR* copies from previously unstudied taxa, we designed primers based on a consensus of known *DFR* copies from the Convolvulaceae. Initially, forward primer DFR59F: 5'-GCGTCACCGGAGCTGCTGG-3' was paired with DFR405R: 5'-GAGGAAGTGAAAACAGC-3' or the GeneRacer oligo dT in 3' rapid amplification of complementary DNA (cDNA) ends (RACE) reactions using either genomic DNA or cDNA synthesized from total RNA as template. Gel-extracted PCR products were cloned into TOPO-TA 2.1 vectors (Invitrogen). Individual clones were sequenced using BigDye version 3.0 (Applied Biosystems) and visualized on ABI 3730 sequencers. To amplify the variable 5' ends of the *DFR* copies, we designed nested gene-specific primers and carried out 5' RACE reactions using the 5' RACE System (Invitrogen) according to the manufacturer's specifications. To confirm all copies from a given taxon had been identified, we also used multiple pairs of degenerate primers in PCRs using independently synthesized cDNA pools as template (primer sequences available on request). Gene-specific primers were then used to amplify the full-length cDNAs of each paralogue, which were cloned into TOPO-TA 2.1 vectors and transformed into Top10 cells (Invitrogen) for storage at -80°C .

Phylogenetic analysis of gene-tree topology, changes in selection and timing of duplication. We combined new sequences collected in the present study with *Ipomoea* sequences available in GenBank and several outgroups to make a single alignment. Because the amino- and carboxy-terminal regions were highly variable, we were unable to discern positional homology with reasonable certainty. These regions (consensus nucleotide positions 13–66 and 1054–1320) were excluded from all subsequent analyses. Alignments are available on request.

The optimal tree and estimates of branch-support were determined using a bayesian Markov chain Monte Carlo (MCMC) approach in the program MrBayes 3.1.2 (ref. 26). In bayesian analyses, we used two models. In the first model, all sites were included in one partition assumed to evolve according to the general time reversible (GTR) with gamma (+G) model. In the second model, one partition contained all first and second codon positions and a second partition contained all third positions, treating each partition independently with its own GTR+G parameters. MCMC analyses were started with flat priors and run for 2 million generations, a burn-in was discarded and consensus trees were determined using MrBayes. At the conclusion of each MCMC run, we estimated that the chains had converged upon a stable set of parameters by calculating the potential scale reduction factor using MrBayes. Optimal topologies were also estimated using maximum parsimony and maximum likelihood criteria (see Supplementary Methods).

We determined if changes in selection followed the origin of duplicate copies of *DFR* using the codeml program in the PAML 3.15 package¹² to perform lineage and clade-specific analyses of K_a/K_s ratios (ω). We first tested a series of nested models that allowed particular branches to have ω independent of

background ratios: the lineage subtending all duplicated *DFR* copies, the lineage subtending the *DFR-B* clade, the lineage subtending the *DFR-A/C* clade, the lineage subtending the *DFR-A* clade and the lineage subtending the *DFR-C* clade. We also tested each of these branches for a signature of positive selection using the branch-sites test of Zhang *et al.*²⁷, and each of the clades arising from these branches using the clade-sites test of Yang *et al.*²⁸. We applied a sequential Bonferroni correction to account for the multiple comparisons made in these analyses and report only results that were significant after this correction (see Supplementary Methods).

In vitro assays of enzyme function. Full-length cDNA transcripts of each *DFR* copy from *I. purpurea*, *C. arvensis*, *E. glomeratus* and *S. lycopersicon* were amplified with Invitrogen Pfx polymerase and cloned directly into pENTR/D-TOPO vectors and then recombined into pDEST17 vectors using Clonase 2.1 (Invitrogen) according to the manufacturer's specifications. Plasmids containing inserts in the correct orientation and coding frame were transformed into BL21 Star (DE3) *E. coli* for overexpression. After overnight growth in 5 ml LB-Amp₁₀₀, we inoculated 200 ml of each culture and grew the cells at 37°C and 200 r.p.m. until mid-log phase ($\text{OD}_{600} = 0.4$). To induce expression, isopropyl- β -D-thiogalactopyranoside was added to a final concentration of 0.5 mM and the cultures were grown for 3 h at 37°C . Cultures were then spun down and the pellet resuspended in lysis buffer (50 mM HEPES, 50 mM NaCl, 1 mM EDTA, 1 mM PMSF, 100 μM TPCK, 100 μM TLCK), lysed for 30 min on ice with 1 mg ml⁻¹ lysozyme and pelleted again leaving the protein-containing supernatant.

Substrates and standards for the enzyme assays were obtained from commercial manufacturers: dihydrokaempferol (TransMit), dihydroquercetin (Alexis Biochemical), dihydromyricetin (Apin Chemicals), eriodictiol (Indofine), naringenin (Sigma), pelargonidin (Indofine), delphinidin (Polyphenols Laboratories AS), cyanidin (Indofine), apigeninidin (Apin) and luteolinidin (Apin). Enzyme extract (250 μl) was added to reaction mixtures for *in vitro* functional enzyme assays as described previously²⁹. For each gene construct, two replicates each of 30 min and 3 h time points were taken. The immediate product of these assays are unstable leucoanthocyanidins (DHK, DHQ and DHM as substrates) or flavan-4-ols (naringenin and eriodictiol as substrates). The former were converted to stable anthocyanidins as described²⁹ and the latter as described³⁰. Products and their appropriate standards were identified spectrophotometrically and by thin-layer chromatography and then quantified spectrophotometrically on a Shimadzu UV-2401PC. Two negative control assays were also run: one assayed overexpressed *I. purpurea* chalcone synthase enzyme, the other substituted lysis buffer for enzyme.

25. Stefanovic, S., Krueger, L. & Olmstead, R. G. Monophyly of the convolvulaceae and circumscription of their major lineages based on DNA sequences of multiple chloroplast loci. *Am. J. Bot.* **89**, 1510–1522 (2002).
26. Ronquist, F. & Huelsenbeck, J. MrBayes 3: bayesian phylogenetic inference under mixed models. *Bioinformatics* **19**, 1572–1574 (2003).
27. Zhang, J., Nielsen, R. & Yang, Z. Evaluation of an improved branch-site likelihood method for detecting positive selection at the molecular level. *Mol. Biol. Evol.* **22**, 2472–2479 (2005).
28. Yang, Z., Wong, W. S. & Nielsen, R. Bayes empirical bayes inference of amino acid sites under positive selection. *Mol. Biol. Evol.* **22**, 1107–1118 (2005).
29. Stafford, H. A. & Lester, H. H. Enzymatic and nonenzymatic reduction of (+)-dihydroquercetin to its 3,4-diol. *Plant Physiol.* **70**, 695–698 (1982).
30. Stafford, H. A. Flavonoids and related phenolic compounds produced in the first internode of *Sorghum vulgare* Pers. in darkness and in light. *Plant Physiol.* **40**, 130–138 (1965).

LETTERS

Genome-scale DNA methylation maps of pluripotent and differentiated cells

Alexander Meissner^{1,2,3*}, Tarjei S. Mikkelsen^{2,4*}, Hongchang Gu², Marius Wernig¹, Jacob Hanna¹, Andrey Sivachenko², Xiaolan Zhang², Bradley E. Bernstein^{2,5,6}, Chad Nusbaum², David B. Jaffe², Andreas Gnirke², Rudolf Jaenisch^{1,7} & Eric S. Lander^{1,2,7,8}

DNA methylation is essential for normal development^{1–3} and has been implicated in many pathologies including cancer^{4,5}. Our knowledge about the genome-wide distribution of DNA methylation, how it changes during cellular differentiation and how it relates to histone methylation and other chromatin modifications in mammals remains limited. Here we report the generation and analysis of genome-scale DNA methylation profiles at nucleotide resolution in mammalian cells. Using high-throughput reduced representation bisulphite sequencing⁶ and single-molecule-based sequencing, we generated DNA methylation maps covering most CpG islands, and a representative sampling of conserved non-coding elements, transposons and other genomic features, for mouse embryonic stem cells, embryonic-stem-cell-derived and primary neural cells, and eight other primary tissues. Several key findings emerge from the data. First, DNA methylation patterns are better correlated with histone methylation patterns than with the underlying genome sequence context. Second, methylation of CpGs are dynamic epigenetic marks that undergo extensive changes during cellular differentiation, particularly in regulatory regions outside of core promoters. Third, analysis of embryonic-stem-cell-derived and primary cells reveals that ‘weak’ CpG islands associated with a specific set of developmentally regulated genes undergo aberrant hypermethylation during extended proliferation *in vitro*, in a pattern reminiscent of that reported in some primary tumours. More generally, the results establish reduced representation bisulphite sequencing as a powerful technology for epigenetic profiling of cell populations relevant to developmental biology, cancer and regenerative medicine.

DNA methylation can be detected by sequencing genomic DNA that has been treated with sodium bisulphite⁷. It has been impractical to apply bisulphite sequencing at a genome-wide scale because polymerase chain reaction (PCR)-based⁸ and whole-genome shotgun⁹ approaches are currently too inefficient for comparative analysis across multiple cell states in large mammalian genomes. However, reduced representations can be generated to sequence a defined fraction of a large genome^{6,10}. Computational analysis indicated that digesting mouse genomic DNA with the methylation-insensitive restriction enzyme MspI, selecting 40–220-base pair (bp) fragments, and performing 36-bp end-sequencing would cover ~1 million distinct CpG dinucleotides (4.8% of all CpGs), with roughly half located within ‘CpG islands’ (including sequences from 90% of all CpG islands) and the rest distributed between other relatively CpG-poor sequence features (Supplementary Fig. 1 and Supplementary Table

1). Notably, although CpGs are not distributed uniformly in the genome, every MspI reduced representation bisulphite sequencing (RRBS) sequence read includes at least one informative CpG position (Supplementary Fig. 2), making the approach highly efficient.

We validated high-throughput RRBS by sequencing MspI fragments from wild-type and methylation-deficient embryonic stem (ES) cells⁶, using an Illumina Genome Analyser. We generated an initial set of ~21 million high quality, aligned RRBS reads. The reads from each cell type included ~97% of the predicted non-repetitive MspI fragments (12-fold and 8-fold median coverage, respectively). This demonstrates that RRBS library construction is relatively unbiased (Supplementary Fig. 3) and is insensitive to genome-wide CpG methylation levels (estimated by nearest-neighbour analysis as 72% and 0.5%, respectively). Reads from both cell types showed near complete (>99%) bisulphite conversion of non-CpG cytosines.

To investigate cell-type-specific DNA methylation patterns, we generated 140 million additional RRBS reads (5.8 gigabase (Gb); Supplementary Information) from ES-derived neural precursor cells (NPCs) and various primary cell populations (Supplementary Table 2). We also generated new chromatin-state maps of H3 lysine 4 mono- and di-methylation (H3K4me1 and H3K4me2) from ES cells, NPCs and whole brain tissue (Supplementary Table 3 and Supplementary Information), using chromatin immunoprecipitation followed by high-throughput sequencing (ChIP-Seq)¹¹.

The methylation levels of CpG dinucleotides in wild-type ES cells display a bimodal distribution (Fig. 1), with most being either ‘largely unmethylated’ (<20% of reads showing methylation) or ‘largely methylated’ (>80% of reads). As expected^{2,8,12}, CpGs in regions of high CpG density (>7% over 300 bp) tend to be unmethylated, whereas CpGs in low-density regions (<5%) tend to be methylated. However, we noted that ~10% of CpGs in low-density regions were unmethylated, whereas ~0.3% of CpGs in high-density regions were methylated. We found that DNA methylation patterns were better explained by histone methylation patterns than by CpG density. Because genomic features tend to be associated with distinct histone methylation patterns¹¹, we analysed these features separately.

High-CpG-density promoters (HCPs) are associated with two classes of genes: ubiquitous ‘housekeeping’ genes and highly regulated ‘key developmental’ genes¹³. In ES cells, HCPs at housekeeping genes are enriched with the transcription initiation mark H3K4me3 (‘univalent’) and are generally highly expressed, whereas those at developmental genes are enriched with both H3K4me3 and the repressive mark H3K27me3 (‘bivalent’) and are generally silent^{11,14}.

¹Whitehead Institute for Biomedical Research, 9 Cambridge Center, Cambridge, Massachusetts 02142, USA. ²Broad Institute of MIT and Harvard, 7 Cambridge Center, Cambridge, Massachusetts 02142, USA. ³Department of Stem Cell and Regenerative Biology, Harvard University, Cambridge, Massachusetts 02138, USA. ⁴Division of Health Sciences and Technology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA. ⁵Molecular Pathology Unit and Center for Cancer Research, MGH, Charlestown, Massachusetts 02129, USA. ⁶Department of Pathology, Harvard Medical School, Boston, Massachusetts 02115, USA. ⁷Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA. ⁸Department of Systems Biology, Harvard Medical School, Boston, Massachusetts 02114, USA.

*These authors contributed equally to this work.

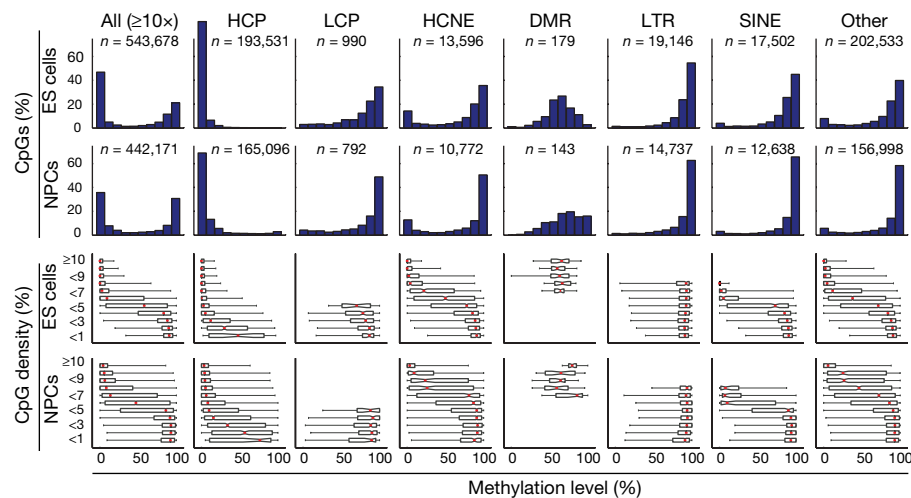


Figure 1 | CpG methylation levels in ES cells and NPCs for CpGs with ≥ 10 -fold coverage. The top histograms show the distribution of methylation levels (%) across all CpGs, HCPs, LCPs, HCNEs, differentially methylated regions (DMRs), LTRs, SINEs and other genomic features (n , number of CpGs). Methylation levels are bimodal (except at DMRs, which have a unimodal distribution largely consistent with uniform sampling from the

maternal and paternal alleles in ES cells and partial hypermethylation in NPCs). The bottom box plots show the distribution of methylation levels conditional on local CpG density (defined as fraction of CpGs in a 300-bp window; shown as percentage). The red lines denote medians, notches the standard errors, boxes the interquartile ranges, and whiskers the 2.5th and 97.5th percentiles.

Both types of promoters are also enriched with H3K4me2, which is associated with an open chromatin confirmation. Out of the 10,299 HCPs sampled (on average, 19 distinct CpGs per promoter), we found that virtually all contain a core region of unmethylated CpGs, regardless of their level of expression or H3K27me3 enrichment (Figs 1 and 2a)^{12,14,15}.

Low-CpG-density promoters (LCPs) are generally associated with tissue-specific genes. In ES cells, a small subset of LCPs are enriched with H3K4me3 ($\sim 7\%$) or H3K4me2 ($\sim 3\%$), and essentially none are enriched with H3K27me3 (ref. 11). We found that whereas most CpGs located in sampled LCPs (990 sites from 392 promoters) are methylated, those in LCPs enriched with H3K4me3 or H3K4me2 have significantly reduced methylation levels (Supplementary Fig. 4).

Distal regulatory regions such as enhancers, silencers and boundary elements are often required to establish correct gene expression patterns in mammalian cells¹⁶. *Cis*-regulatory elements active in a particular cell type are often associated with markers of open chromatin, such as H3K4me2 or H3K4me1 (refs 17, 18). We identified 25,051 sites of H3K4me2 enrichment in ES cells from 1 kb to >100 kb

away from known promoters (most were also enriched with H3K4me1, but not with H3K4me3). CpGs sampled at H3K4me2-enriched sites (outside of promoters and CpG islands) had significantly lower methylation levels than those at unenriched sites (Fig. 2b). This relationship was particularly strong for CpGs located in highly conserved non-coding elements (HCNEs; Fig. 2c).

Imprinting control regions (ICRs) are CpG-rich regulatory regions that display allele-specific histone and DNA methylation¹⁹. Our RRBS library included sequences from 13 of ~ 20 known ICRs (on average, 13 distinct CpGs per ICR). CpGs within these elements display a unimodal distribution of methylation levels, with a median close to 50%, which is consistent with hypomethylation of the active allele marked with H3K4me3 and hypermethylation of the silenced allele marked with H3K9me3 (Fig. 1)¹¹.

Interspersed repeat families differ in their chromatin structure, with H3K9me3 enriched at active long terminal repeats (LTRs) and to a lesser extent at long interspersed elements (LINEs), but not at short interspersed elements (SINEs). Notably, CpGs located in LTRs and LINEs are generally hypermethylated, even in CpG-rich contexts

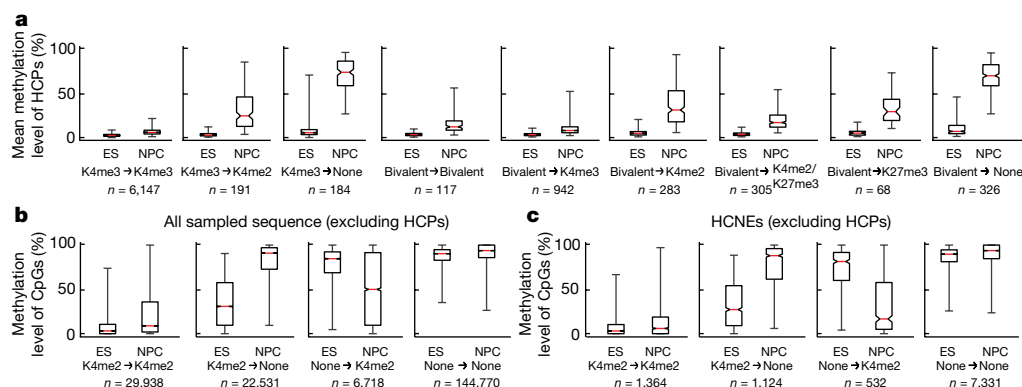


Figure 2 | Correlation between DNA and histone methylation. **a**, Mean methylation levels across CpGs within each profiled HCP (requiring ≥ 5 -fold coverage of ≥ 5 CpGs), conditional on their histone methylation state in ES cells and NPCs (n , number of HCPs; those enriched with H3K4me3 are generally also enriched for H3K4me2, but not vice versa). Loss of H3K4 methylation, and to a lesser extent of H3K27me3, is correlated with gain of DNA methylation. **b**, Methylation levels of individual CpGs outside of HCPs, conditional on enrichment of H3K4me2 (n , number of distinct sites

in each category). Changes in histone methylation state are inversely correlated with changes in DNA methylation. **c**, Methylation levels of CpGs in HCNEs not overlapping CpG islands, conditional on H3K4me2 enrichment. For **a–c**, the red lines denote medians, notches the standard errors, boxes the interquartile ranges, and whiskers the 2.5th and 97.5th percentiles. All pair-wise comparisons of methylation levels at sites with changing chromatin states are significant ($P < 10^{-20}$, Mann–Whitney U test).

(Fig. 1). In contrast, CpGs in SINEs show a correlation between methylation levels and CpG density that is comparable to non-repetitive sequences.

We conclude that in ES cells the presence of H3K4 methylation and the absence of H3K9 methylation are better predictors of unmethylated CpGs than sequence context alone. This is consistent with models in which *de novo* methyl-transferases either specifically recognize sites with unmethylated H3K4 (ref. 20) or are excluded by H3K4 methylation or associated factors. Similarly, H3K9me3 or associated factors may recruit methyl-transferases at ICRs and repetitive elements²¹.

We next used RRBS to analyse how DNA methylation patterns change when ES cells are differentiated *in vitro* into a homogeneous population of NPCs (Supplementary Fig. 4)²². Whereas CpG methylation levels are highly correlated between the two cell types ($\rho = 0.81$), there were clear differences: ~8% of CpGs unmethylated in ES cells became largely methylated in NPCs, whereas ~2% of CpGs methylated in ES cells became unmethylated; these changes were strongly correlated with changes in histone methylation patterns.

At both univalent and bivalent HCPs, we found that most CpGs remained unmethylated on differentiation, particularly within their core CpG island, but that loss of H3K4me3 and retention of H3K4me2 or H3K27me3 correlated with a partial increase in DNA methylation levels (median, ~25%; 2.9% and 32% of univalent and bivalent HCPs, respectively) and complete loss of H3K4 and H3K27 methylation correlated with DNA hypermethylation (median, ~75%; 2.8% and 16% of univalent and bivalent HCPs, respectively; Fig. 2).

Most LCPs marked by H3K4 methylation in ES cells lose this mark in NPCs; however, LCPs associated with genes expressed in NPCs gain this mark. Loss or gain of H3K4 methylation is a strong predictor of inverse changes in CpG methylation levels at these promoters (Supplementary Fig. 5).

Our chromatin-state maps revealed that 18,899 (75%) of putative distal regulatory elements enriched with H3K4me2 in ES cells lost this mark in NPCs, whereas 20,088 new H3K4me2 sites appeared, often in HCNE-rich regions surrounding activated developmental genes (Fig. 3). Loss or gain of H3K4 methylation were again inversely correlated with CpG methylation levels (Fig. 2b, c). In fact, these regions account for most observed de-methylation events. The presence of H3K27me3 alone did not correlate with lower methylation levels in CpG-poor regions (Supplementary Fig. 6).

The data support the notion that CpG-rich and -poor regulatory elements undergo distinct modes of epigenetic regulation^{2,11,12}. Most (>95%) HCPs seem to be constitutively unmethylated and regulated by trithorax-group (trxG; associated with H3K4me3) and/or Polycomb-group (PcG; associated with H3K27me3) proteins, which may be recruited in part by means of non-specific unmethylated-CpG binding domains²³. Hypermethylation of these CpG-dense regions leads to exclusion of trxG/PcG activity, heterochromatin formation and essentially irreversible gene silencing². In contrast, regulatory elements in CpG-poor sequence contexts seem to undergo extensive and dynamic methylation and de-methylation. Hence, methylation of isolated CpGs may contribute to chromatin condensation or directly interfere with transcription factor binding², but does not necessarily prevent chromatin remodelling in response to activating signals.

As noted above, a small set of HCPs ($n = 252$; ~3%) became hypermethylated (>75% mean methylation across sampled CpGs) on *in vitro* differentiation of ES cells to NPCs. To investigate whether the observed pattern reflects an *in vivo* regulatory mechanism, we isolated NPCs from embryonic day (E)13.5 embryos and differentiated them into glial fibrillary acidic protein (Gfap)-positive astrocytes (with no more than two passages *in vitro*). We similarly differentiated the *in vitro*-derived NPCs into astrocytes (with these cells having undergone at least 18 passages; Supplementary Fig. 4), and compared the two populations using RRBS (Fig. 4a–f).

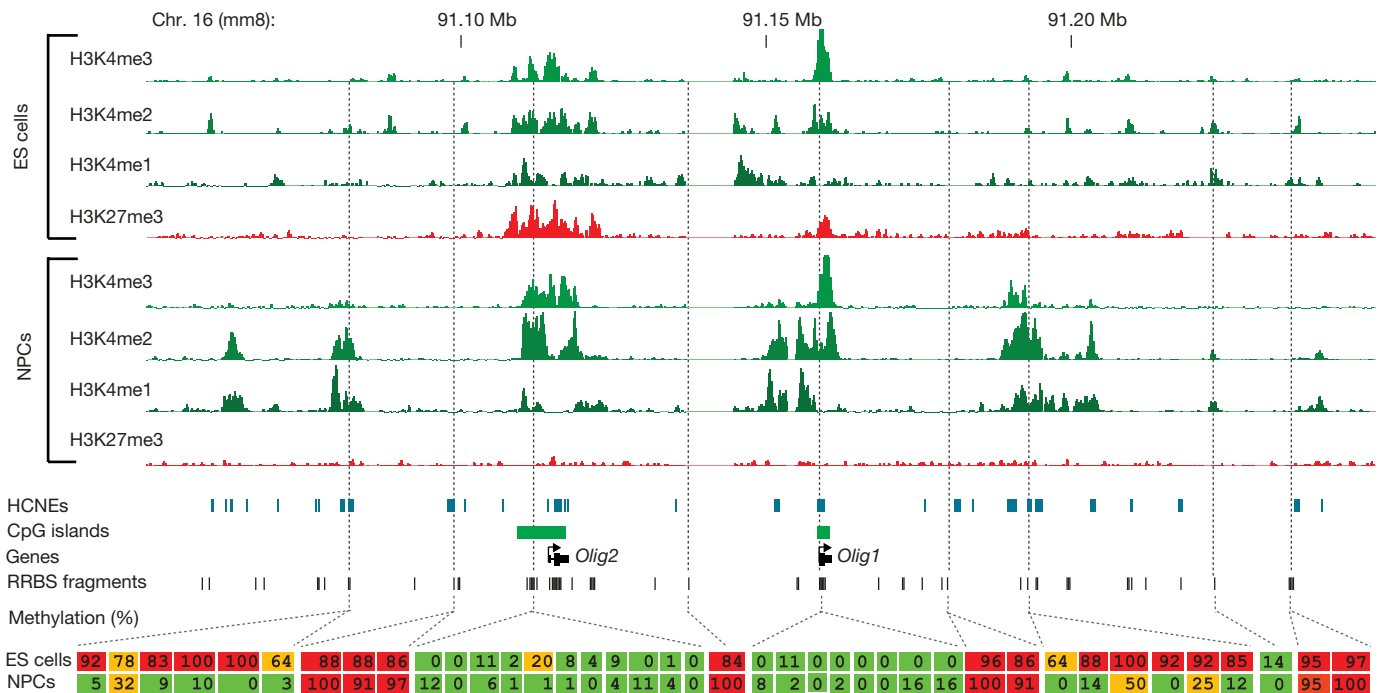


Figure 3 | Developmentally regulated de-methylation of highly conserved non-coding elements. Comparison of histone and DNA methylation levels across the *Olig1/Olig2* neural-lineage transcription factor locus. ChIP-Seq tracks for H3K4me1/2/3 and H3K27me3 in ES cells and NPCs are shown. The unmethylated CpG-rich promoters are bivalent and inactive in ES cells and resolve to univalent H3K4me3 on activation in NPCs. H3K4me2

enrichment appears over HCNEs distal to the two genes, and this correlates with CpG de-methylation. Inferred methylation levels for 40 out of 215 sampled CpGs are shown and colour-coded. Red indicates largely methylated (>80%); green indicates largely unmethylated (<20%), and orange indicates intermediate levels (≥20% and ≤80%).

The methylation levels of CpGs were highly correlated ($\rho = 0.85$), but astrocytes obtained from *in vivo* NPCs displayed substantially less HCP hypermethylation than those obtained from ES cells (Fig. 4a). The *in vivo*-derived astrocytes showed hypermethylation at only 30 HCPs, largely associated with germline-specific genes (including *Dazl*, *Hormad1*, *Sycp1*, *Sycp2* and *Taf7l*), several of which also showed partial methylation in ES cells. In contrast, the *in vitro*-derived astrocytes showed hypermethylation of these and ~305 additional HCPs. This set includes some genes known to be expressed by at least some *in vivo* astrocytes (including *Isyn1*, *Gsn* and *Cldn5*; ref. 24) but that were silent in the ES-cell-derived astrocytes (Supplementary Information). However, the hypermethylated HCPs are significantly enriched for genes not expressed in NPCs or in the astrocyte lineage (Supplementary Tables 4–7). They include genes involved in development and differentiation of neuronal (*Lhx8*, *Lhx9*, *Moxd1*, *Htr1f* and *Slit1*), ependymal (*Otx2* and *Kl*) and unrelated lineages (including *Myod1*, *Dhh* and *Nkx3-1*). In fact, we found that ‘key developmental’ HCPs that are bivalent in ES cells are six times more likely to be included in the hypermethylated set compared to univalent HCPs. Moreover, univalent genes in the hypermethylated set are expressed at significantly lower levels in both ES cells and primary astrocytes, compared to those that remained hypomethylated (Fig. 4g). We also found that the hypermethylated HCPs tend to have a ~15% lower CpG density (Fig. 4h).

To investigate further the differences between *in vitro* and *in vivo* cell populations, we analysed whole brain tissue (representing cells of mainly glial origin). Virtually all (>99%) of sampled HCPs were

unmethylated (Fig. 4c) and enriched with H3K4me3 and/or H3K27me3 (Supplementary Fig. 7), with ~20 germline-specific HCPs being the only clear exceptions. RRBS libraries from other *in vivo* sources (T cells, B cells, spleen, lung, liver and fibroblasts) also showed few hypermethylated HCPs (Supplementary Fig. 8). This suggests that—apart from silencing germline-specific¹², imprinted and X-inactivated (Supplementary Fig. 9) genes in somatic tissues—hypermethylation of HCPs is not a major mechanism of developmental regulation *in vivo*.

To test for a correlation between passage number and HCP hypermethylation, we examined independently derived *in vitro* NPCs collected after only 9 passages. These cells displayed hypermethylation at approximately half of the HCPs that are hypermethylated in the NPCs after 18 passages (Fig. 4d, e). To reduce time in culture further, we used Sox1–GFP (green fluorescent protein) ES cells²⁵ to isolate very early NPCs. These cells initially displayed virtually no HCP hypermethylation. However, after continued culturing they acquired hypermethylation at many of the same HCPs as the previous NPC populations (Supplementary Fig. 8). Finally, we grew the *in vivo*-derived NPCs for 11 passages *in vitro*, differentiated them into astrocytes and then examined the methylation pattern. Notably, these cells had also begun to acquire hypermethylation at a largely similar set of HCPs (Fig. 4a, b).

These results show that independently derived NPC populations from both *in vitro* and *in vivo* sources and different genetic backgrounds reproducibly undergo gradual hypermethylation at a characteristic set of HCPs. These observations have several implications.

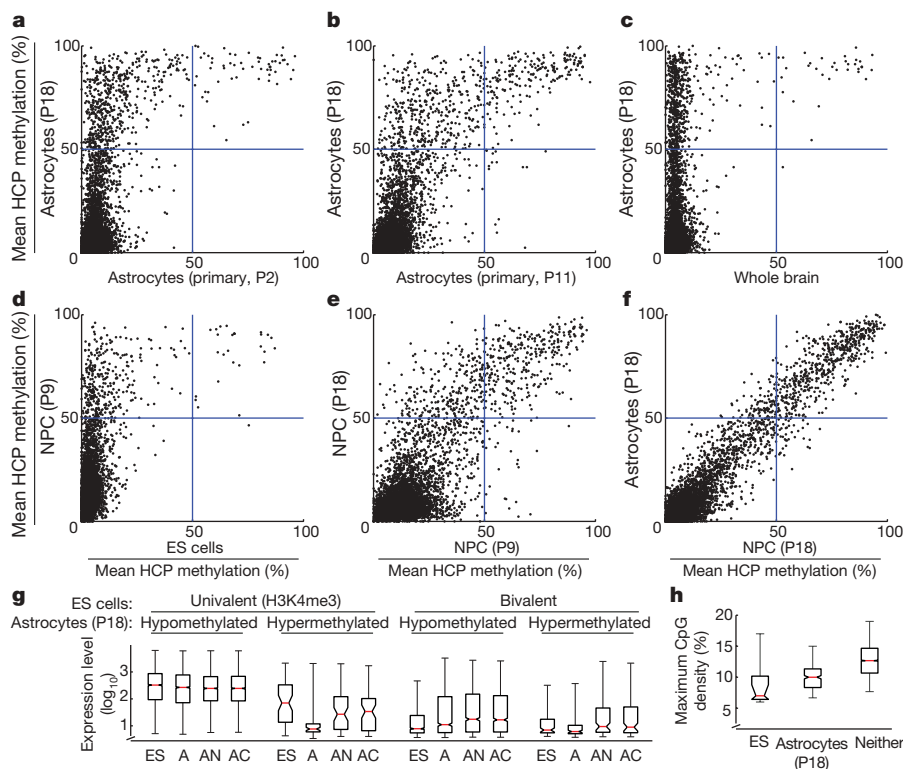


Figure 4 | HCP hypermethylation of cultured cells. Inferred mean methylation levels (%) across autosomal HCPs (requiring ≥ 5 -fold coverage of ≥ 5 CpGs within the CpG island). **a**, ES-derived astrocytes contains roughly 10 times more hypermethylated HCPs than primary NPC-derived astrocytes after two passages (P) in culture. **b**, Continued passage of the primary cells lead to gradual hypermethylation of many of the same HCPs. **c**, Only a handful of mainly germline-specific HCPs display hypermethylation in a whole brain tissue sample. **d**, Most HCPs are unmethylated in ES cells, but a small subset gain significant methylation on differentiation to NPCs. **e**, Continued proliferation of NPCs leads to additional HCPs becoming hypermethylated after 18 passages. **f**, Differentiation of late-stage NPCs into astrocytes by growth factor

withdrawal does not lead to additional HCP hypermethylation. **g**, Expression levels of genes associated with profiled HCPs for ES cells (ES), ES-derived astrocytes (A), primary neocortical astrocytes (AN) and cerebellar astrocytes (AC). Hypermethylation of HCPs is correlated with low expression levels in ES-derived astrocytes. HCPs that are univalent in ES cells and become hypermethylated in ES-derived astrocytes are associated with lower expression levels in both ES cells and primary astrocytes. **h**, The maximal CpG densities (300-bp window) of hypermethylated HCPs in ES cells or ES-derived astrocytes are significantly lower than for unmethylated HCPs. For **g** and **h**, the red lines denote medians, notches the standard errors, boxes the interquartile ranges, and whiskers the 2.5th and 97.5th percentiles.

First, aberrant epigenetic regulation in culture has raised concern over the accuracy of cellular models generated by *in vitro* differentiation or manipulation^{26–28}. Both primary and transformed cell lines, including ES-derived NPC populations, tend to lose developmental potency after continued proliferation in culture^{26,29}. Susceptibility to hypermethylation at key regulatory genes that are normally activated on differentiation could explain this phenomenon. Second, malignant cells are often found to harbour hypermethylated CpG islands^{4,5}. Recently, genes known to undergo frequent hypermethylation in adult cancers were noted to be significantly enriched for genes with bivalent promoters in ES cells (reviewed in ref. 30). The similarities between hypermethylation in culture and in cancer may provide a useful *in vitro* model for studying a common underlying mechanism. Finally, the gradual hypermethylation of ‘weak’ HCPs hints at underlying kinetics. Because H3K4 methylases are targeted, at least in part, by non-specific CpG-binding domains²³, such HCPs may be particularly sensitive to imbalanced chromatin-modifying factors or other cancer- or culture-related perturbations.

More generally, RRBS makes it feasible to perform genome-scale bisulphite sequencing on large-mammalian genomes, providing a valuable tool for epigenetic profiling of cell populations. As sequencing capacity increases, genome coverage can be readily scaled in step by adding restriction enzymes, increasing the selected size range or using hybridization-based reduced representation strategies.

METHODS SUMMARY

ES cells and ES-derived neural cells were cultured as described previously^{11,25}. Primary tissues were isolated from 4–6-week-old male 129SvJae/C57/B6 mice. Mouse embryonic fibroblasts (MEFs) and primary neural precursors were isolated from 129SvJae/C57/B6 E14.5 embryos.

RRBS libraries were prepared from 1–10 µg mouse genomic DNA digested with 10–100 Units MspI (NEB). Size-selected MspI fragments (40–120 bp and 120–220 bp) were filled in and 3′-terminal-A extended, extracted with phenol and precipitated with ethanol. Ligation to pre-annealed adapters containing 5′-methyl-cytosine instead of cytosine (Illumina) was performed using the Illumina DNA preparation kit and protocol. QIAquick (Qiagen) cleaned-up, adaptor-ligated fragments were bisulphite-treated using the EpiTect Bisulphite Kit (Qiagen). Preparative-scale PCR was performed and QIAquick-purified PCR products were subjected to a final size selection on a 4% NuSieve 3:1 agarose gel. SYBR-green-stained gel slices containing adaptor-ligated fragments of 130–210 bp or 210–310 bp in size were excised. Library material was recovered from the gel (QIAquick) and sequenced on an Illumina 1G genome analyser.

Sequence reads from bisulphite-treated Solexa libraries were identified using standard Illumina base-calling software and then analysed using a custom computational pipeline. ChIP-Seq experiments, sequencing, alignments and identification of significantly enriched regions were carried out as described previously¹¹.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 24 March; accepted 21 May 2008.

Published online 6 July 2008.

- Bestor, T. H. The DNA methyltransferases of mammals. *Hum. Mol. Genet.* **9**, 2395–2402 (2000).
- Bird, A. DNA methylation patterns and epigenetic memory. *Genes Dev.* **16**, 6–21 (2002).
- Reik, W. Stability and flexibility of epigenetic gene regulation in mammalian development. *Nature* **447**, 425–432 (2007).
- Feinberg, A. P. The epigenetics of cancer etiology. *Semin. Cancer Biol.* **14**, 427–432 (2004).
- Jones, P. A. & Baylin, S. B. The epigenomics of cancer. *Cell* **128**, 683–692 (2007).
- Meissner, A. *et al.* Reduced representation bisulfite sequencing for comparative high-resolution DNA methylation analysis. *Nucleic Acids Res.* **33**, 5868–5877 (2005).

- Frommer, M. *et al.* A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. *Proc. Natl Acad. Sci. USA* **89**, 1827–1831 (1992).
- Eckhardt, F. *et al.* DNA methylation profiling of human chromosomes 6, 20 and 22. *Nature Genet.* **38**, 1378–1385 (2006).
- Cokus, S. J. *et al.* Shotgun bisulphite sequencing of the *Arabidopsis* genome reveals DNA methylation patterning. *Nature* **452**, 215–219 (2008).
- Altshuler, D. *et al.* An SNP map of the human genome generated by reduced representation shotgun sequencing. *Nature* **407**, 513–516 (2000).
- Mikkelsen, T. S. *et al.* Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* **448**, 553–560 (2007).
- Weber, M. *et al.* Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. *Nature Genet.* **39**, 457–466 (2007).
- Saxonov, S., Berg, P. & Brutlag, D. L. A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters. *Proc. Natl Acad. Sci. USA* **103**, 1412–1417 (2006).
- Bernstein, B. *et al.* A bivalent chromatin structure marks key Developmental genes in embryonic stem cells. *Cell* **125**, 315–326 (2006).
- Illingworth, R. *et al.* A novel CpG island set identifies tissue-specific methylation at developmental gene loci. *PLoS Biol.* **6**, e22 (2008).
- West, A. G. & Fraser, P. Remote control of gene transcription. *Hum. Mol. Genet.* **14** (Spec No 1) R101–R111 (2005).
- Heintzman, N. D. *et al.* Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nature Genet.* **39**, 311–318 (2007).
- Bernstein, B. E. *et al.* Genomic maps and comparative analysis of histone modifications in human and mouse. *Cell* **120**, 169–181 (2005).
- Edwards, C. A. & Ferguson-Smith, A. C. Mechanisms regulating imprinted genes in clusters. *Curr. Opin. Cell Biol.* **19**, 281–289 (2007).
- Ooi, S. K. *et al.* DNMT3L connects unmethylated lysine 4 of histone H3 to *de novo* methylation of DNA. *Nature* **448**, 714–717 (2007).
- Esteve, P. O. *et al.* Direct interaction between DNMT1 and G9a coordinates DNA and histone methylation during replication. *Genes Dev.* **20**, 3089–3103 (2006).
- Conti, L. *et al.* Niche-independent symmetrical self-renewal of a mammalian tissue stem cell. *PLoS Biol.* **3**, e283 (2005).
- Voo, K. S., Carlone, D. L., Jacobsen, B. M., Flodin, A. & Skalik, D. G. Cloning of a mammalian transcriptional activator that binds unmethylated CpG motifs and shares a CXXC domain with DNA methyltransferase, human trithorax, and methyl-CpG binding domain protein 1. *Mol. Cell. Biol.* **20**, 2108–2121 (2000).
- Sharma, M. K. *et al.* Distinct genetic signatures among pilocytic astrocytomas relate to their brain region origin. *Cancer Res.* **67**, 890–900 (2007).
- Aubert, J. *et al.* Screening for mammalian neural genes via fluorescence-activated cell sorter purification of neural precursors from *Sox1-gfp* knock-in mice. *Proc. Natl Acad. Sci. USA* **100** (Suppl 1), 11836–11841 (2003).
- Jones, P. A., Wolkowicz, M. J., Harrington, M. A. & Gonzales, F. Methylation and expression of the Myo D1 determination gene. *Phil. Trans. R. Soc. Lond. B* **326**, 277–284 (1990).
- Smiraglia, D. J. *et al.* Excessive CpG island hypermethylation in cancer cell lines versus primary human malignancies. *Hum. Mol. Genet.* **10**, 1413–1419 (2001).
- Shen, Y., Chow, J., Wang, Z. & Fan, G. Abnormal CpG island methylation occurs during *in vitro* differentiation of human embryonic stem cells. *Hum. Mol. Genet.* **15**, 2623–2635 (2006).
- Bouhous, I. A., Joannides, A., Kato, H., Chandran, S. & Allen, N. D. Embryonic stem cell-derived neural progenitors display temporal restriction to neural patterning. *Stem Cells* **24**, 1908–1913 (2006).
- Ohm, J. E. & Baylin, S. B. Stem cell chromatin patterns: an instructive mechanism for DNA hypermethylation? *Cell Cycle* **6**, 1040–1043 (2007).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank the staff of the Broad Institute Genome Sequencing Platform for assistance with data generation and B. Ramsahoye for the nearest neighbour analysis. This research was supported by funds from the National Human Genome Research Institute, the National Cancer Institute, and the Broad Institute of MIT and Harvard.

Author Information All primary sequencing data have been submitted to the NCBI GEO repository under accession numbers GSE11034 (RRBS), GSE11172 (ChIP-Seq) and GSE11483 (gene expression microarrays). Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to R.J. (jaenisch@wi.mit.edu) or E.S.L. (lander@broad.mit.edu).

METHODS

Cell culture and ES cell differentiation. V6.5 (129/B6), Sox1-EGFP knock-in (Sox1-GFP 129/129)²⁵ and methylation-deficient (Dnmt1^{kd}, 3a^{-/-}, 3b^{-/-}) ES cells were expanded on γ -irradiated MEFs in DMEM plus 15% fetal bovine serum (FBS, Hyclone) supplemented with 1 \times MEM-nonessential amino acids (Life Technologies), 0.1 mM 2-mercaptoethanol and 10³ Units ml⁻¹ leukaemia inhibitory factor (LIF). After passaging onto gelatin-coated dishes (0.1% gelatin, Sigma), ES cells were trypsinized and transferred to bacterial dishes allowing embryoid body formation. Embryoid bodies were propagated for 4 days in the same medium in the absence of LIF and subsequently plated onto tissue culture dishes. One day after plating, the medium was replaced by ITSFn; that is, DMEM/F12 (Life Technologies) supplemented with 5 μ g ml⁻¹ insulin, 50 μ g ml⁻¹ human APO transferrin, 30 nM sodium selenite (all Sigma), 2.5 μ g ml⁻¹ fibronectin and penicillin/streptomycin (both Life Technologies). After 5–7 days, cells were trypsinized, triturated to a single cell suspension, replated on laminin-coated dishes (1 μ g ml⁻¹, Life Technologies) and further propagated in N3 medium composed of DMEM/F12, 25 μ g ml⁻¹ insulin, 50 μ g ml⁻¹ transferrin, 30 nM sodium selenite, 20 nM progesterone, 100 nM putrescine (Sigma), 10 ng ml⁻¹ Fgf2 (R&D Systems) and penicillin/streptomycin. Neural precursor cell proliferation was maintained by daily additions of Fgf2. Sox1-EGFP-positive neural precursors were isolated and FACS-purified (FACS Aria, Becton Dickinson) either from ITSFn cultures or after short-term expansion in Fgf2. Growth factor withdrawal of these cultures results in terminal differentiation into primarily neuronal cell populations. Neural precursor cell lines were obtained by sequential passaging and propagation in the presence of 20 ng ml⁻¹ Egf and 10 ng ml⁻¹ Fgf2 (both R&D Systems). Differentiation into astrocytes was induced by growth factor withdrawal and addition of 5% FBS for 5 days.

Primary tissues and cell types. Primary tissues were isolated from 4–6-week-old male 129SvJae/C57/B6 mice. MEFs and primary neural precursors were isolated from 129SvJae/C57/B6 E14.5 embryos. MEFs were generated according to standard protocols. *In vivo* neural precursors were isolated by disaggregating the whole brain and plating the suspension under the conditions described previously. Established lines were differentiated into astrocytes by growth factor withdrawal and addition of serum (see previously).

MspI RRBS library construction. 1–10 μ g mouse genomic DNA was digested with 10–100 Units of MspI (NEB) in a 30–500 μ l reaction 16–20 h at 37 °C. Digested DNA was extracted with phenol, precipitated with ethanol and size-selected on a 4% NuSieve 3:1 agarose gel (Lonza). DNA marker lanes were excised from the gel and stained with SYBR Green (Invitrogen). For each sample, two slices containing DNA fragments of approximately 40–120 bp and 120–220 bp, respectively, were excised from the unstained preparative portion of the gel. DNA was recovered using Easy Clean DNA spin filters (Primm Labs), extracted with phenol and precipitated with ethanol. The two size fractions were kept apart throughout the procedure, including during the final sequencing. Size-selected MspI fragments were filled in and 3'-terminal A extended in a 50 μ l reaction containing 20 U Klenow exo⁻ (NEB), 0.4 mM dATP, 0.04 mM dGTP and 0.04 mM 5-methyl-dCTP (Roche) in 1 \times NEB buffer 2 (15 min at 25 °C followed by 15 min at 37 °C), extracted with phenol and precipitated with ethanol using 10 μ g glycogen (Roche) as a carrier. Ligation to pre-annealed Illumina adapters containing 5'-methyl-cytosine instead of cytosine (Illumina) was performed using the Illumina DNA preparation kit and protocol. QIAquick (Qiagen) cleaned-up, adaptor-ligated fragments were bisulphite-treated using the EpiTect Bisulphite Kit (Qiagen) with minor modifications:

the bisulphite conversion time was increased to approximately 14 h by adding three cycles (5 min of denaturation at 95 °C followed by 3 h at 60 °C). After bisulphite conversion, the single-stranded uracil-containing DNA was eluted in 20 μ l of elution buffer. Analytical (25 μ l) PCR reactions containing 0.5 μ l of bisulphite-treated DNA, 5 pmol each of genomic PCR primers 1.1 and 2.1 (Illumina) and 2.5 U PfuTurboC α Hotstart DNA polymerase (Stratagene) were set up to determine the minimum number of PCR cycles required to recover enough material for sequencing. Preparative-scale (8 \times 25 μ l) PCR was performed using the same PCR profile: 5 min at 95 °C, $n\times$ (30 s at 95 °C, 20 s at 65 °C, 30 s at 72 °C), followed by 7 min at 72 °C, with n ranging from 18 to 24 cycles. QIAquick-purified PCR products were subjected to a final size selection on a 4% NuSieve 3:1 agarose gel. SYBR-green-stained gel slices containing adaptor-ligated fragments of 130–210 bp or 210–310 bp in size were excised. RRBS library material was recovered from the gel (QIAquick) and sequenced on an Illumina 1G genome analyser.

Sequence alignments and data analysis. Sequence reads from bisulphite-treated Solexa libraries were identified using standard Illumina base-calling software and then analysed using a custom computational pipeline. Residual cytosines (Cs) in each read were first converted to thymines (Ts), with each such conversion noted for subsequent analysis. A reference sequence database was constructed from the 36-bp ends of each computationally predicted MspI fragment in the 40–220-bp size range. All Cs in each fragment end were then converted to Ts (only the C-poor strands are sequenced in the RRBS process; Supplementary Fig. 2).

The converted reads were aligned to the converted reference by finding all 12-bp perfect matches and then extending to both ends of the treated read, not allowing gaps (reverse complement alignments were not considered). The number of mismatches in the induced alignment was then counted between the unconverted read and reference, ignoring cases in which a T in the unconverted read is matched to a C in the unconverted reference. For a given read, the best alignment was kept if the second-best alignment had ≥ 2 more mismatches, otherwise the read was discarded as non-unique. Low-quality reads were identified and discarded if $\sum_{q \in Q} 10^{q/10} > 1,000$, where Q denotes the read quality scores at each mismatched position. The methylation level of each sampled cytosine was estimated as the number of reads reporting a C, divided by the total number of reads reporting a C or T, counting only reads with quality scores of ≥ 20 at the position.

HCP, ICP and LCP annotations were taken from ref. 11. CpG island and other annotations were downloaded from the UCSC browser (mm8). Estimation of methylation levels from individual CpGs was limited to those with ≥ 10 -fold coverage. The methylation level of an HCP promoter was estimated as the mean methylation level across all CpGs with ≥ 5 -fold coverage overlapping the annotated CpG island(s) in the promoter, requiring at least 5 such CpGs. HCPs were classified as hypermethylated if this mean methylation level was $\geq 75\%$.

Chromatin immunoprecipitation. H3K4me1 (ab8895), H3K4me2 (ab7766) and H3K4me3 (ab8580) antibodies were purchased from Abcam. ChIP experiments on mouse ES cells (H3K4me1/2), NPCs (H3K4me1/2) and whole brain tissue (H3K4me1/2/3), Illumina/Solexa sequencing, alignments and identification of significantly enriched regions (using 1 kb sliding windows and correction for alignability) were carried out as described previously¹¹.

Expression data. RNA expression data for ES-derived astrocytes were generated as described previously¹¹ and analysed using GenePattern (<http://www.broad.mit.edu/cancer/software/genepattern/>). Primary astrocyte data were obtained from ref. 24.

cAMP signalling in mushroom bodies modulates temperature preference behaviour in *Drosophila*

Sung-Tae Hong¹, Sunhoe Bang¹, Seogang Hyun¹, Jongkyun Kang¹, Kyunghwa Jeong¹, Donggi Paik¹, Jongkyeong Chung¹ & Jaeseob Kim¹

Homoiotherms, for example mammals, regulate their body temperature with physiological responses such as a change of metabolic rate and sweating. In contrast, the body temperature of poikilotherms, for example *Drosophila*, is the result of heat exchange with the surrounding environment as a result of the large ratio of surface area to volume of their bodies^{1,2}. Accordingly, these animals must instinctively move to places with an environmental temperature as close as possible to their genetically determined desired temperature. The temperature that *Drosophila* instinctively prefers has a function equivalent to the 'set point' temperature in mammals. Although various temperature-gated TRP channels have been discovered^{3,4}, molecular and cellular components in *Drosophila* brain responsible for determining the desired temperature remain unknown. We identified these components by performing a large-scale genetic screen of temperature preference behaviour (TPB) in *Drosophila*^{5,6}. In parallel, we mapped areas of the *Drosophila* brain controlling TPB by targeted inactivation of neurons with tetanus toxin⁷ and a potassium channel (Kir2.1)⁸ driven with various brain-specific GAL4s. Here we show that mushroom bodies (MBs) and the cyclic AMP–cAMP-dependent protein kinase A (cAMP–PKA) pathway are essential for controlling TPB. Furthermore, targeted expression of cAMP–PKA pathway components in only the MB was sufficient to rescue abnormal TPB of the corresponding mutants. Preferred temperatures were affected by the level of cAMP and PKA activity in the MBs in various PKA pathway mutants.

To identify genes controlling TPB in *Drosophila*, TPBs of over 27,000 independent P-element mutants were examined^{5,6}. The avoidance index against low temperatures (AI_{low}) and avoidance index against high temperatures (AI_{high}) were calculated from the following formulae:

$$AI_{low} = \frac{N_{int} - N_{low}}{N_{int} + N_{low}}$$

and

$$AI_{high} = \frac{N_{int} - N_{high}}{N_{int} + N_{high}}$$

where N_{low} is the number of flies in the region below 21.5 °C, N_{high} is the number of flies in the region above 26.5 °C, and N_{int} is the number of flies in the intermediate region, between 21.5 and 26.5 °C (Supplementary Fig. 1). From the genetic screen we identified multiple genes involved in MB formation or learning and memory processes. The genes were *oskar*, *pka C1*, *amnesiac*, *14-3-3ζ*, *dCREB2*, *ptpmeg*, *mud*, *dCBP*, *dunce*, *PKA anchor protein* and *mub* (Supplementary Figs 2 and 3). TPB assay results of previously characterized mutants of these genes and revertants of mutants we isolated confirmed that these genes are indeed involved in TPB control

(Supplementary Figs 2 and 4, Supplementary Results and Discussion). These results led us to hypothesize that MBs and cAMP–PKA signalling might be key components of TPB control.

In parallel with the TPB genetic screens, we mapped brain regions that control TPB. Discrete parts of the fly brain were impaired by targeted expression of tetanus toxin (TNT) with various brain-specific GAL4 genes (Fig. 1a, Supplementary Fig. 5 and Supplementary Table 2). As shown in Fig. 1a, *GAL4* genes driving TNT in MBs, dorsal paired medial (DPM) neurons and antenna lobes resulted in a loss of the preference for 24–25 °C. DPM neurons release AMN neuropeptide to MBs and modulate learning and memory^{9,10}. Antenna lobes receive signals from the antenna, an organ known to sense temperature¹¹. These results are consistent with our genetic screen findings and support the hypothesis that the MB is a key cellular component for TPB.

To exclude the possibility of developmental defects induced by TNT, we combined *UAS–TNT* with the Temporal and Regional Gene Expression Targeting (TARGET) system⁹, which employs the gene encoding the temperature-sensitive GAL80 protein (*GAL80^{ts}*) under the control of *tubulin* promoter (*tub–GAL80^{ts}*). In this system, transcriptional activator activity of GAL4 is suppressed at 18 °C but is active at 32 °C. The *MB247>UAS–TNT; tub–GAL80^{ts}* flies showed normal TPB at 18 °C. However, flies raised at 18 °C until they were three-day-old adults and then exposed for 16 h at 32 °C just before the TPB assay showed abnormal TPB (Fig. 1c), which is similar to *MB247>UAS–TNT*. When the *MB247>UAS–TNT; tub–GAL80^{ts}* flies were transferred back to 18 °C for 3 days of recovery after the 16-h heat shock at 32 °C, they showed completely normal TPB (Fig. 1c, Supplementary Fig. 6 and Supplementary Table 3). These data show that the alteration of TPB by TNT in MBs is reversible and is not the result of neuronal death or developmental defects.

Because TNT blocks only neural activity requiring *n-Syb*-dependent synaptic vesicle fusion¹², we ectopically expressed *UAS–Kir2.1* with the *GAL4* genes (Fig. 1b). Overexpression of Kir2.1 hyperpolarizes neurons and suppresses action-potential firing^{8,12}. Although many *GAL4* genes driving Kir2.1 were lethal, *189y* and *17d GAL4* flies having *UAS–Kir2.1* were viable and showed abnormal TPB similar to that of the *GAL4* flies with *UAS–TNT* (Fig. 1b). To avoid lethality, we combined TARGET with *UAS–Kir2.1*. As shown in Fig. 1d, all three MB-specific *GAL4* flies tested were extremely cryophilic when Kir2.1 was induced just before the TPB assay. This phenotype was very similar to flies in which MBs were ablated with hydroxyurea (Fig. 1e and Supplementary Fig. 7e). Kir2.1 expression in the DPM with *c316* also resulted in cryophilic behaviour (Fig. 1b).

TNT inactivates only TNT-sensitive chemical synapses but not electrical or TNT-insensitive chemical synapses^{12,13}. In contrast, Kir2.1 silences all types of synapse by hyperpolarization¹². In our

¹Department of Biological Sciences, Korea Advanced Institute of Science & Technology, Guseong-Dong, Yuseong-Gu, Daejeon 305-701, Korea.

experiments, complete inactivation of MBs with Kir2.1 or complete removal of MBs with hydroxyurea resulted in extreme preference for cold (Fig. 1d, e). In contrast, selective inactivation of TNT-sensitive synapses among MB neurons resulted in a perturbation of the ability to find a preferred temperature (Fig. 2a–c and Supplementary Fig. 14). They were spread widely over the temperature gradient. These TPBs were similar to the TPBs of *mbm*¹ (MB miniature)¹⁴ and *ceb*¹ (central brain deranged)¹⁴ mutants (Fig. 1f, Supplementary Fig. 7 and Supplementary Table 4), which have partial defects in MB neurons. These data imply that coordinated functions of the neurons constituting MBs are critical in enabling flies to find the desired temperature.

We also examined whether other areas of the fly brain are involved in TPB control (Fig. 1a, b). We found that none of ellipsoid body (EB), fan-shaped body (FB), neurons of the proboscis and mouthparts, subesophageal ganglia (SOG), lateral protocerebrum cells, pars intercerebralis (PI), corpora cardiaca (CC) and core clock neurons were required for TPB control (see Supplementary Information).

The MBs consist of three groups of neurons ($\alpha\beta$, $\alpha'\beta'$ and γ lobes), and each lobe is functionally distinct (Supplementary Fig. 10)^{9,15,16}. By expressing TNT and Kir2.1 with TARGET, we investigated which lobe is involved in TPB control (Supplementary Table 5). As shown in Fig. 2, all GAL4s active in $\alpha\beta$ and $\alpha'\beta'$ lobes resulted in the dispersal of flies over the temperature gradient after heat shock at 32 °C (Fig. 2a–c). In contrast, GAL4s active only in the γ lobe showed almost normal TPB (Fig. 2d–f). Taken together, these data show that the $\alpha\beta$ lobe of the MB is essential for proper TPB, and the γ lobe is dispensable.

To examine the role of cAMP–PKA signalling in TPB, we examined previously known mutants of the cAMP–PKA pathway. *DC0* trans-heterozygotes of *X4/tw2*, *X4/ γ 15*, *X4/MB581* and *B10/tw2*, which are hypomorphic mutants of the catalytic subunit of PKA, were widely dispersed over the temperature gradient (Fig. 3a, Supplementary Fig. 11 and Supplementary Table 6).

Next we tested genetic interactions between upstream components of PKA. Heterozygotes for *amn*, *rut* and *DC0* showed normal TPB

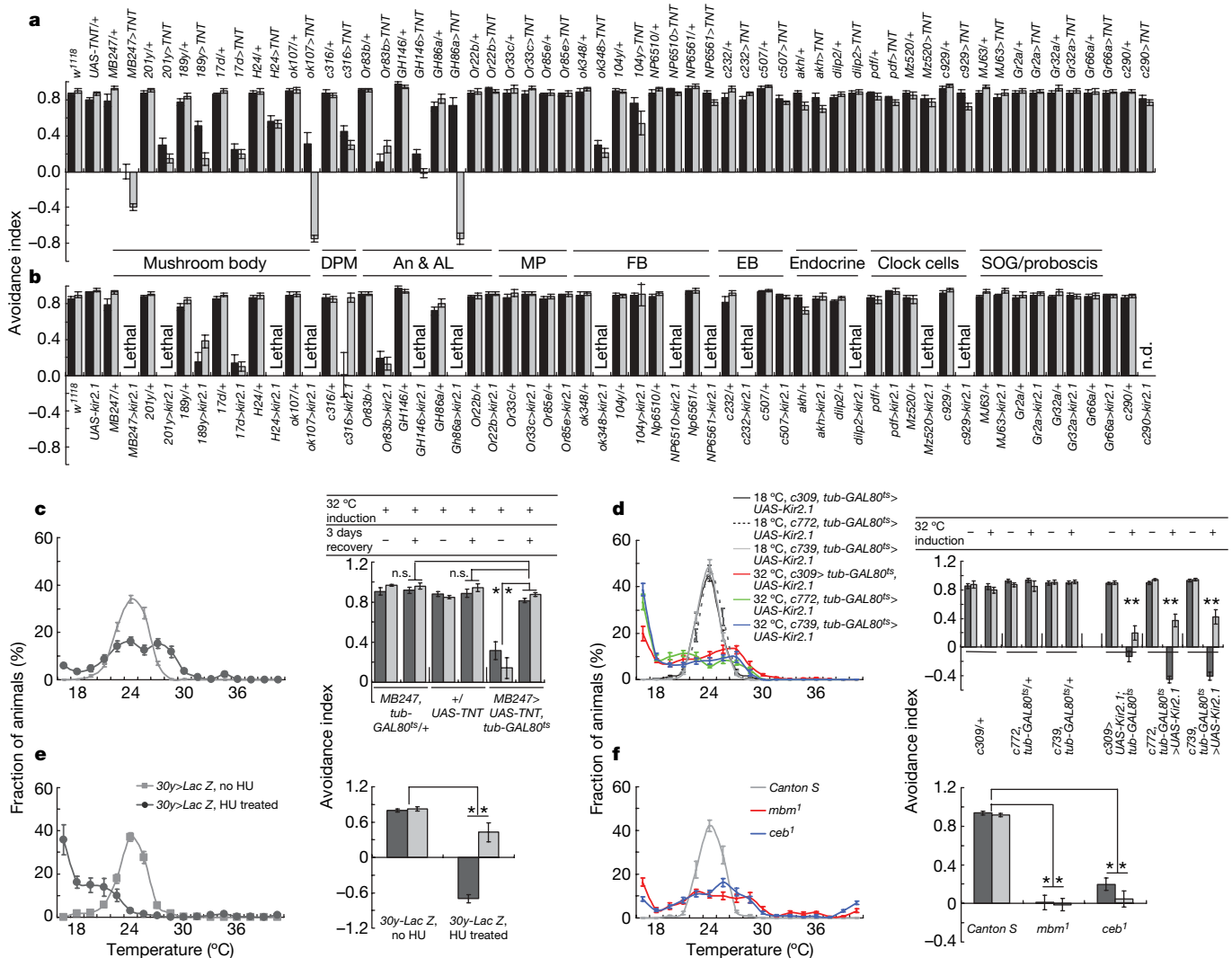


Figure 1 | The MB is required for TPB. **a, b**, The effect of brain-specific GAL4 genes driving TNT (**a**) or Kir2.1 (**b**) on TPB. Neural activity in fly brains was inhibited by using GAL4-driven UAS–TNT or UAS–Kir2.1, and the resulting AI indices were calculated. Thirteen GAL4 lines carrying Kir2.1 showed developmental lethality (Lethal). DPN, dorsal paired neurons; An & AL, antennae and antennal lobe; MP, maxillary palp; FB, fan-shaped body; EB, ellipsoid body; SOG, subesophageal ganglion; n.d., not determined. Dark bars, AI_{low}; grey bars, AI_{high}. **c**, *MB247>UAS-TNT; tub-GAL80ts*. Conditional TNT expression was induced by exposure to 32 °C for 16 h (black line). The effect of TNT on MBs was reversible. TNT induced for 16 h

by 32 °C heat shock was followed by recovery for 3 days at 18 °C (grey line). **d**, Temporal induction of Kir2.1 in MBs with the TARGET system just before TPB assay resulted in cryophilic TPB. All flies were raised at 18 °C. For Kir2.1 induction, flies were incubated at 32 °C for 16 h just before TPB assay. **e**, The effect of MB ablation by hydroxyurea (HU) on TPB. *30y>UAS-LacZ* flies were treated with hydroxyurea to ablate the MB. The MB-ablated flies showed impaired temperature preference. **f**, Genetic mutants with abnormal MB morphologies. *mbm*¹ and *ceb*¹ showed abnormal TPB. Two asterisks, *P* < 0.001 by ANOVA. All data are means and s.e.m.

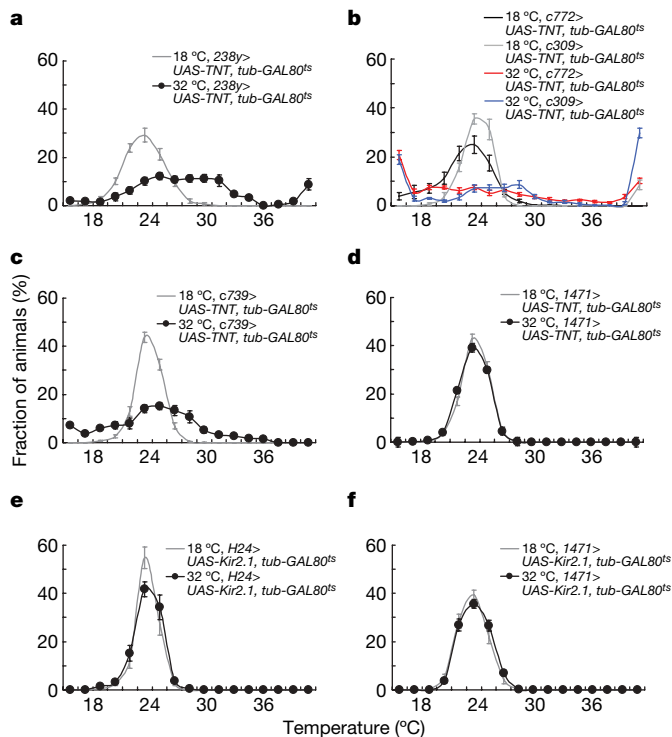


Figure 2 | Contribution of MB lobes in TPB control. Neural activity of each MB lobe was conditionally blocked in adult flies. **a**, $\alpha\beta$, $\alpha'\beta'$ and γ lobe neurons blocked by $238y>UAS-TNT$; $tub-GAL80^{ts}$. **b**, $\alpha\beta$ and γ lobe neurons blocked by $c772>UAS-TNT$; $tub-GAL80^{ts}$ and $c309>UAS-TNT$; $tub-GAL80^{ts}$. **c**, $\alpha\beta$ lobe neurons blocked by $c739>UAS-TNT$; $tub-GAL80^{ts}$. **d**, γ lobe neurons blocked by $1471>UAS-TNT$; $tub-GAL80^{ts}$. **e**, γ lobe neurons blocked by $UAS-Kir2.1$, conditionally. $H24>UAS-Kir2.1$; $tub-GAL80^{ts}$ (**e**); $1471>UAS-Kir2.1$; $tub-GAL80^{ts}$ (**f**). All data are means \pm s.e.m.

(Supplementary Figs 11b and 12a, b). However, double heterozygotes of *amn*/*rut*, *amn*/*DC0* and *rut*/*DC0* showed abnormal TPB and a notable decrease in both AI_{low} and AI_{high} (Fig. 3b and Supplementary Fig. 13). These data strongly support cAMP–PKA signalling as a critical factor in TPB.

rutabaga and *dunce* encode adenylyl cyclase and cAMP phosphodiesterase, respectively. The *rut¹* mutant, in which the cAMP level is decreased¹⁷, had a decreased avoidance of low temperature, whereas *dnc¹/dnc¹* and *dnc¹/dnc¹M14*, in which the cAMP level is elevated¹⁸, showed a decreased avoidance of high temperature (Fig. 3c). This raises the possibility that flies with low levels of cAMP prefer a lower temperature and flies with high levels of cAMP prefer a higher temperature. This idea was confirmed by experiments with *UAS-PKA^{CA}* (encoding a constitutively active form of PKA) and *UAS-PKA^{DN}* (encoding a dominant-negative form of PKA). After 24 h of feeding with RU486, *elav-GS(GeneSwitch)>UAS-PKA^{CA}* flies showed a decreased avoidance of high temperature but not low temperature, whereas the *elav-GS>UAS-PKA^{DN}* flies fed with RU486 showed a marked decrease in low-temperature avoidance but relatively little decrease in high-temperature avoidance (Fig. 3d). Overall, these results demonstrate that the temperatures that flies prefer are correlated with the level of cAMP and with PKA activity.

This correlation was clearer when we restricted ectopic PKA^{CA} and PKA^{DN} to MBs in the adult stage with TARGET (Fig. 3e–h). After exposure to 32 °C for 16 h, both *UAS-PKA^{DN}* flies with MB247; *tub-GAL80^{ts}* and *c772*; *tub-GAL80^{ts}* showed markedly less avoidance of low temperature (Fig. 3e, f). The shift of preferred temperature was more marked in *PKA^{CA}*. After heat shock at 32 °C for 16 h, the axes of distribution profiles of MB247; *tub-GAL80^{ts}*>*UAS-PKA^{CA}* and *c772*; *tub-GAL80^{ts}*>*UAS-PKA^{CA}* flies were shifted from 24 °C to 27 °C (Fig. 3g, h). Collectively, these results imply that the level of cAMP

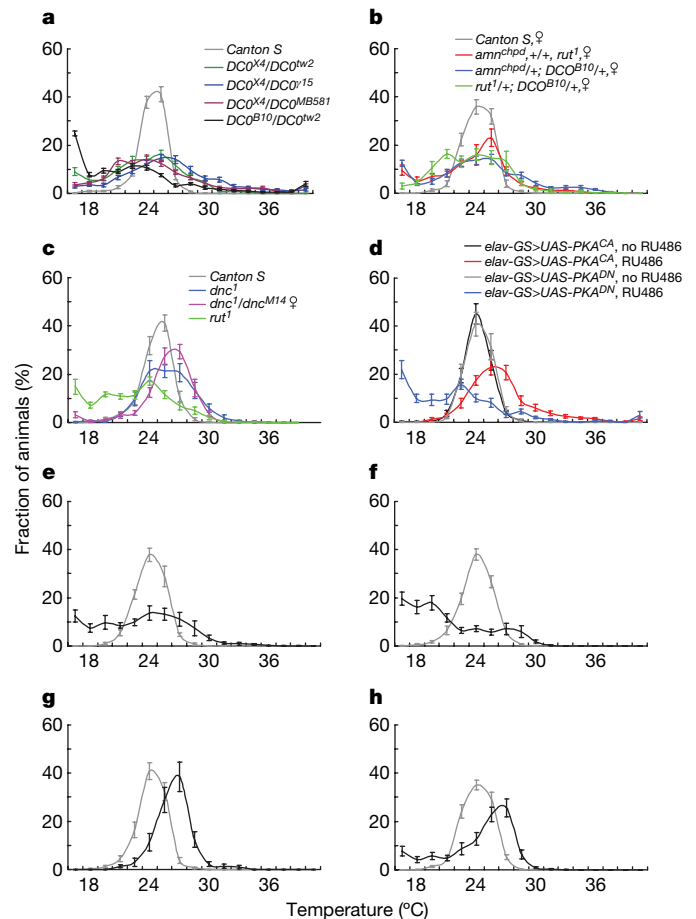


Figure 3 | cAMP–PKA signalling in the MB is needed for normal temperature preference. **a**, Impaired PKA activity caused abnormal temperature preference. *DC0* trans-heterozygotes showed considerably dispersed distribution profiles compared with the wild type. **b**, Genetic interaction between cAMP signalling mutants. Double mutants for *amn*, *rut* and *DC0* showed significantly altered TPB. **c**, Temperature preference of *dnc* and *rut* mutants. Distribution profiles of *dnc* mutants showed decreased avoidance of high temperature. *rut¹* mutants showed decreased avoidance of low temperature. **d**, The pan-neuronal driver *elav-GeneSwitch* (GS) was used to express *UAS-PKA^{CA}* (constitutively active) or *UAS-PKA^{DN}* (dominant-negative) conditionally. RU486 (100 μ g ml⁻¹) was administered for 24 h at the adult stage. **e–h**, MB-restricted impairment of PKA activity shifts the preferred temperature: MB247; *tub-GAL80^{ts}*>*UAS-PKA^{DN}* (**e**), *c772*; *tub-GAL80^{ts}*>*UAS-PKA^{DN}* (**f**), MB247; *tub-GAL80^{ts}*>*UAS-PKA^{CA}* (**g**), *c772*; *tub-GAL80^{ts}*>*UAS-PKA^{CA}* (**h**). Suppressed PKA transgene expression at 18 °C (grey lines) was released by incubation for 16 h at 32 °C (black lines). All data are means \pm s.e.m.

and the activity of PKA within the MB is a critical factor in determining the temperature preferred by the flies. Finding the right temperature seems to be the main defect of flies with low cAMP/PKA, whereas the temperature preferred by flies with high cAMP/PKA was simply shifted higher (Fig. 3). This suggests that cAMP/PKA serves two distinct functions: recognizing temperature and determining the desired temperature.

These results led us to question whether cAMP–PKA signalling controls TPB through the MB alone or to the MB and other parts of the fly body. We generated flies in which normal cAMP–PKA signalling is maintained only in the MB but not elsewhere. This was achieved by MB-restricted ectopic expression of the *rutabaga* and *PKA* transgenes with MB247 or *c309* GAL4 in the background of *rut¹* and *DC0* (Fig. 4a–e). The *rut¹* mutants carrying MB247>*UAS-rut* or *c309>UAS-rut* showed almost completely restored TPBs (Fig. 4a, b). We also ectopically expressed PKA only in MBs of *DC0^{MB581}/DC0^{X4}* and *DC0^{tw2}/DC0^{X4}* with MB247 or *c309*

GAL4 genes. These flies also showed almost completely restored TPBs (Fig. 4c–e). These data indicate that cAMP/PKA signalling in the MB is necessary and sufficient for normal TPB, whereas normal levels of cAMP/PKA signalling in other parts of the body are dispensable.

Our results raise the question of whether temperature sensation and interpretation require learning and memory (LM) processes. Our data do not exclude this possibility, but there are some differences between

LM and TPB. Targeted expression of TNT in the MB does not affect LM¹² but alters TPB. The γ lobe of the MB is necessary for LM^{19–21} but not for TPB (Fig. 2). Moreover, we observed that mutants of *NFI* and *TbH*, which are important genes in LM⁹, do not show any TPB defects. This implies that temperature sensation and interpretation might share molecular and cellular mechanisms with learning and memory, but they do not use identical processes.

METHODS SUMMARY

Fly strains and MB ablation. Fly stocks were raised on standard cornmeal food at 25 °C and 40–50% relative humidity. *Canton-S* (C-S), *w¹¹¹⁸* and *yw* were used as wild-type *Drosophila* strains in this study. We generated *UAS-TNT*; *tub-GAL80^{ts}* flies by crossing *UAS-TNT* (ref. 5) and *tub-GAL80^{ts}* (Stock 7018; Bloomington) flies. *UAS-TNT*; *UAS-GFP* flies were generated by crossing *UAS-TNT* and *UAS-GFP* flies. Both *c772* carrying *tub-GAL80^{ts}* and MB247 carrying *tub-GAL80^{ts}* were described previously¹⁹. To ablate MB cells, a paste of heat-killed yeast with or without 50 mg ml⁻¹ hydroxyurea was fed to newly hatched first-instar larvae (*30y-GAL4/UAS-LacZ*) for 4 h as described elsewhere^{22,23}.

Behavioural screen and TPB assay. Behavioural screening for TPB was performed as described previously^{5,6}. A linear temperature gradient from 15 to 45 °C was established along the aluminium block. Flies were incubated under a 12 h:12 h dark:light cycle for four days before TPB assays. All TPB assays were performed in a room with 35 ± 5% relative humidity. The $\Delta T^{3,4,6,24}$ for 21.5–26.5 °C bins (intermediate temperature region) versus $T^{\circ}\text{C}$ (low-temperature or high-temperature region) was calculated as described in the text. Flies were aged for three to six days in a 12 h:12 h dark:light cycle before temperature preference was assayed. Unless stated otherwise, a mixed population of sexes was tested for TPB. The number of flies was recorded and processed with Microsoft Excel, and analysis of variance (ANOVA) tests were performed with the SAS general linear model.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 4 February; accepted 15 May 2008.

Published online 29 June 2008; corrected 7 August 2008 (details online).

1. Bear, M. F., Connors, B. W. & Paradiso, M. A. Neuroscience: exploring the brain (Lippincott Williams & Wilkins, Baltimore, MD, 2001).
2. Zars, T. Two thermosensors in *Drosophila* have different behavioral functions. *J. Comp. Physiol. A* **187**, 235–242 (2001).
3. Liu, L., Yermolaieva, O., Johnson, W. A., Abboud, F. M. & Welsh, M. J. Identification and function of thermosensory neurons in *Drosophila* larvae. *Nature Neurosci.* **6**, 267–273 (2003).
4. Rosenzweig, M. et al. The *Drosophila* ortholog of vertebrate *TRPA1* regulates thermotaxis. *Genes Dev.* **19**, 419–424 (2005).
5. Lee, Y. et al. Pyrexia is a new thermal transient receptor potential channel endowing tolerance to high temperatures in *Drosophila melanogaster*. *Nature Genet.* **37**, 305–310 (2005).
6. Hong, S. T. et al. Histamine and its receptors modulate temperature-preference behaviors in *Drosophila*. *J. Neurosci.* **26**, 7245–7256 (2006).
7. Sweeney, S. T., Broadie, K., Keane, J., Niemann, H. & O'Kane, C. J. Targeted expression of tetanus toxin light chain in *Drosophila* specifically eliminates synaptic transmission and causes behavioral defects. *Neuron* **14**, 341–351 (1995).
8. Baines, R. A., Uhler, J. P., Thompson, A., Sweeney, S. T. & Bate, M. Altered electrical properties in *Drosophila* neurons developing without synaptic transmission. *J. Neurosci.* **21**, 1523–1531 (2001).
9. Keene, A. C. & Waddell, S. *Drosophila* olfactory memory: single genes to complex neural circuits. *Nature Rev. Neurosci.* **8**, 341–354 (2007).
10. Krashes, M. J., Keene, A. C., Leung, B., Armstrong, J. D. & Waddell, S. Sequential use of mushroom body neuron subsets during *Drosophila* odor memory processing. *Neuron* **53**, 103–115 (2007).
11. Sayeed, O. & Benzer, S. Behavioral genetics of thermosensation and hygrosensation in *Drosophila*. *Proc. Natl Acad. Sci. USA* **93**, 6079–6084 (1996).
12. Thum, A. S. et al. Differential potencies of effector genes in adult *Drosophila*. *J. Comp. Neurol.* **498**, 194–203 (2006).
13. Rister, J. & Heisenberg, M. Distinct functions of neuronal synaptobrevin in developing and mature fly photoreceptors. *J. Neurobiol.* **66**, 1271–1284 (2006).
14. de Belle, J. S. & Heisenberg, M. Expression of *Drosophila* mushroom body mutations in alternative genetic backgrounds: a case study of the *mushroom body miniature* gene (*mbm*). *Proc. Natl Acad. Sci. USA* **93**, 9875–9880 (1996).
15. Davis, R. L. Olfactory memory formation in *Drosophila*: from molecular to systems neuroscience. *Annu. Rev. Neurosci.* **28**, 275–302 (2005).
16. Skoulakis, E. M. & Grammenoudi, S. Dunces and da Vinci: the genetics of learning and memory in *Drosophila*. *Cell. Mol. Life Sci.* **63**, 975–988 (2006).
17. Levin, L. R. et al. The *Drosophila* learning and memory gene *rutabaga* encodes a Ca^{2+} /calmodulin-responsive adenylyl cyclase. *Cell* **68**, 479–489 (1992).

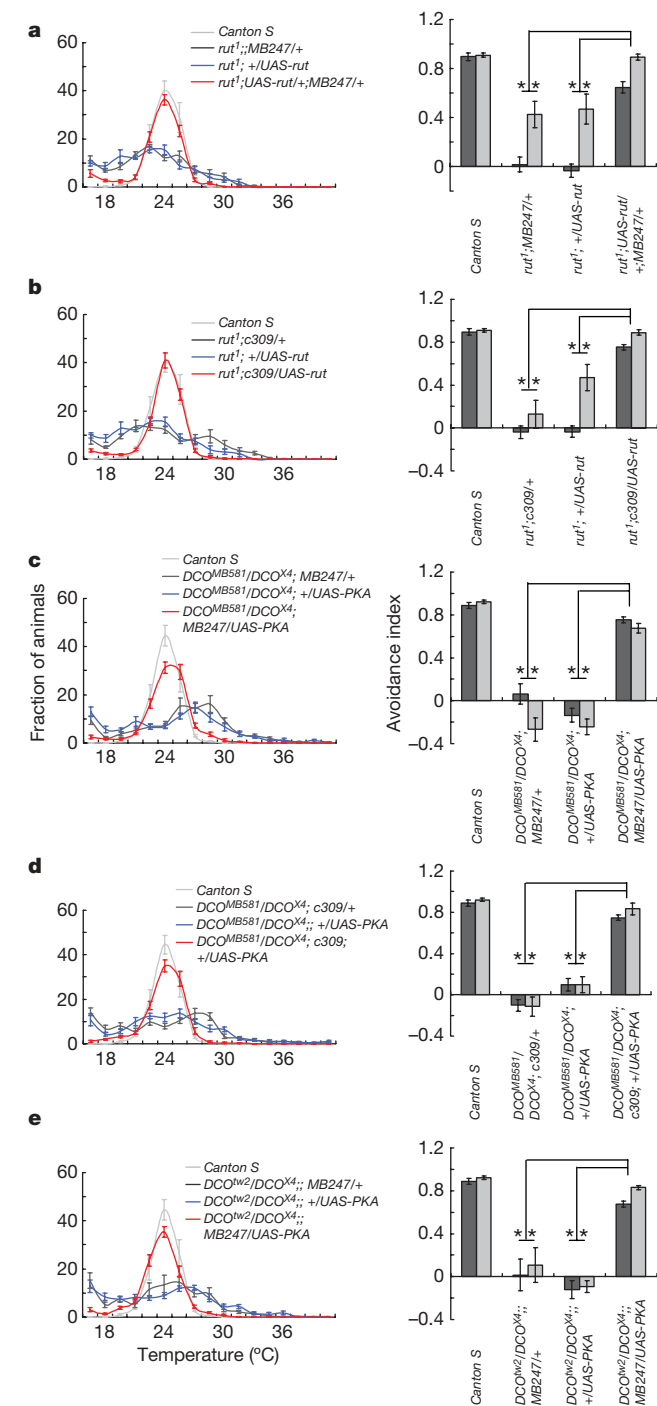


Figure 4 | cAMP-PKA in the MB is sufficient for normal TPBs. **a, b,** MB-restricted *UAS-rut* expression in *rut¹* mutant. **a**, Abnormal temperature preference in *rut¹* mutant was rescued by MB247-driven *rut* transgene. **b**, A *rut¹* mutant carrying *c309>UAS-rut* was also rescued. **c–e**, MB-restricted *UAS-PKA* (catalytic subunit) expression in *DCO^{MB581}/DCO^{X4}* or *DCO^{w2}/DCO^{X4}* mutant. *DCO^{MB581}/DCO^{X4}* flies with MB247>*UAS-PKA* (**c**) or with *c309>UAS-PKA* (**d**) showed repaired temperature preference. **e**, *DCO^{w2}/DCO^{X4}* with MB247>*UAS-PKA*. Two asterisks, $P < 0.001$ by ANOVA. All data are means and s.e.m.

18. Davis, R. L. & Kiger, J. A. Jr. Dunce mutants of *Drosophila melanogaster*: mutants defective in the cyclic AMP phosphodiesterase enzyme system. *J. Cell Biol.* **90**, 101–107 (1981).
19. Ferris, J., Ge, H., Liu, L. & Roman, G. G_o signaling is required for *Drosophila* associative learning. *Nature Neurosci.* **9**, 1036–1040 (2006).
20. Zars, T., Fischer, M., Schulz, R. & Heisenberg, M. Localization of a short-term memory in *Drosophila*. *Science* **288**, 672–675 (2000).
21. Isabel, G., Pascual, A. & Preat, T. Exclusive consolidated memory phases in *Drosophila*. *Science* **304**, 1024–1027 (2004).
22. Rodan, A. R., Kiger, J. A. Jr & Heberlein, U. Functional dissection of neuroanatomical loci regulating ethanol sensitivity in *Drosophila*. *J. Neurosci.* **22**, 9490–9501 (2002).
23. McBride, S. M. *et al.* Mushroom body ablation impairs short-term memory and long-term memory of courtship conditioning in *Drosophila melanogaster*. *Neuron* **24**, 967–977 (1999).
24. Gray, J. M. *et al.* Oxygen sensation and social feeding mediated by a *C. elegans* guanylate cyclase homologue. *Nature* **430**, 317–322 (2004).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank D. Stafford for help in manuscript preparation. This work was supported by grants from the Brain Research Center of the 21st Century Frontier Program funded by the Korean Ministry of Science and Technology and a grant from the Science Research Center (SRC) for Functional Cellulomics of the Korea Science and Engineering Foundation (KOSEF).

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to J.K. (kjaeseob@kaist.ac.kr).

METHODS

Drosophila strains and transgenic lines. *UAS-reaper (rpr)* was provided by J. R. Nambu; other lines were provided as follows: *MB247*, *c772*, *201y*, *189y*, *30y*, *17d*, *H24* (T. Zars), *ok107* (Kei Ito), *c316* (S. Waddell), *Or83b* (G. Miesenböck), *GH86a*, *GH146* (R. F. Stocker), *Or33c*, *Or85e* (J. R. Carlson), *Gr2a*, *Gr32a*, *Gr66a* (K. Scott), *NP6510*, *NP6561* (L. Liu), *c232*, *c739*, *c309*, *ok348* (L. C. Griffith), *c507*, *104y*, *MJ63*, *238y*, *UAS-Kir2.1* (A. Sehgal), *UAS-Kir2.1/CyO*; *tub-GAL80^{ts}* (H. Tanimoto), *akh* (S. K. Kim), *dilp2* (E. Hafen), *pdf* (J. H. Park), *Mz520*, *c929* (F. Rouyer), *c290* (U. Heberlein), *1471* (Bloomington stock centre), *MB247; tub-GAL80^{ts}*, *c772*; *tub-GAL80^{ts}*, *c739*; *tub-GAL80^{ts}* (G. Roman), *DC0* alleles and *UAS-PKA^{CA}* and *PKA^{DN}* (D. Kalderon), *ceb¹*, *mbm¹*, *cbd^{1KS92}*, *ccb^{KS145}*, *ccd^{KS135}*, *ebo⁴* (M. Heisenberg), *osk⁶*, *leo^{2,3}*, *s162*, *ptpmeg¹*, *Df(3L)ED201*, *mud¹*, *mud⁴*, *dnc^{M14}*, *dnc¹* and *rut¹* (Bloomington stock centre), *oc¹* (J.-R. Martin), *MB247-RFP* (A. Fiala), *amn^{chpd}* (U. Heberlein) and *elav-GS* (H. Keshishian).

Relative quantitative RT-PCR. Flies 3–6 days old were collected and frozen immediately in liquid nitrogen. Heads were removed from bodies by vortex-mixing and separated in a chilled sieve. RNA was extracted from heads with an RT-PCR miniprep kit (Stratagene). cDNA was prepared with a reverse transcription system (Promega) with the addition of a random-sequence hexamer. cDNA was analysed by relative quantitative RT-PCR (Ambion). 18S ribosomal RNA was used as an internal control. Data processing was performed with Image J software (Windows version of NIH Image) and Microsoft Excel. To quantify PCR products, the band intensity of the gene-specific amplicon was divided by the intensity of the 18S amplicon.

Conditional induction of the TNT transgene. Continuous expression of *TNT* can affect normal developmental processes²⁵. To minimize these, a *UAS-TNT*; *tub-GAL80^{ts}* fly was prepared. *GAL80^{ts}* is a temperature-sensitive allele for *GAL80* (ref. 26). This can suppress DNA binding of GAL4 protein at 18 °C but not at 32 °C. To induce conditional TNT expression in specific brain regions, we crossed *UAS-TNT*; *tub-GAL80^{ts}* fly with a brain-specific *GAL4* line. These flies were incubated and transferred periodically to new culture bottles at 18 °C. Eclosed F₁ progenies with both a *GAL4* driver and *UAS-TNT*; *tub-GAL80^{ts}* were collected and aged for 3 days at 18 °C before the temperature preference assay. For TNT induction, F₁ progenies were transferred from 18 °C to 32 °C and incubated for 12–16 h. Heterozygous *GAL4* flies and heterozygous *UAS-TNT*; *tub-GAL80^{ts}* flies were used as controls. Breeding temperature and motor activity did not affect temperature preference.

PKA assay. Kinase activity was measured with the Pierce IQ serine/threonine kinase assay kit. Detailed procedures were conducted in accordance with the manufacturer's instructions. In brief, a head extract was made by homogenization of 50 heads in 200 µl of extraction buffer (10 mM sodium phosphate pH 6.7, 1 mM EDTA, 0.5 mM EGTA, 2.5 mM 2-mercaptoethanol, 1 × Roche protease

inhibitor cocktail (Complete mini) and centrifuged at 10,000g for 10 min. Extracted supernatant was saved and its concentration was determined with the Bio-Rad reagent (Bradford). Head extract (22 µl), diluted in enzyme dilution buffer (20 mM HEPES pH 7.4, 1 mM dithiothreitol, 0.05% Triton X-100), was assayed in a total volume of 30 µl containing 50 mM MOPS pH 7.0, 10 mM MgCl₂, 0.25 mg ml⁻¹ BSA, 0.1 mM ATP, 10 µM cAMP and 60 µM peptide substrate (Pierce) at 25 °C. Fluorescence intensity was measured with a λ_{ex} = 560/λ_{em} = 590 filter set. Relative PKA enzyme activity was calculated as follows: Relative enzyme activity is defined as (relative fluorescence units (RFU) of the observed sample minus RFU of the no-enzyme reference sample)/(RFU of control sample (Canton S) minus RFU of the no-enzyme reference sample).

Immunohistochemistry. Adult brains expressing transgenic *UAS-GFP*, *MB247::RFP* (ref. 27) or *UAS-LacZ* were removed from the head capsules and fixed in 4% paraformaldehyde in PBS for 20–30 min, and rinsed in PBS containing 0.5% Triton X-100. Brains with *UAS-LacZ* were incubated with anti-β-galactosidase (1:1,000 dilution; Promega). They were then incubated with the appropriate green-fluorescent-protein labelled secondary antibody (Jackson Laboratories). For staining with anti-FASII (1:75 dilution; Hybridoma Bank), undissected adult heads of the genetic mutants were fixed for 2 h in 4% paraformaldehyde in PBS containing 0.5% Triton X-100. After being washed three times each for 45 min, the fixed heads were dissected. Confocal analysis was performed on a Zeiss LSM5 microscope. Confocal stacks were processed with LSM5 image browser and Adobe Photoshop.

Behavioural screen and TPB assay. Behavioural screening for TPB was performed as described previously^{3,4}. The AI^{1,2,4,24} of 21.5–26.5 °C bins was calculated as described in the text. Flies were aged for 3–6 days in a 12 h:12 h light:dark cycle before temperature preference was assayed. Unless stated otherwise, a mixed population of sexes was tested for TPB. The number of flies was recorded and processed with Microsoft Excel, and ANOVA tests were performed with the SAS general linear model. Sensitivity to environmental effects within each genotype was evaluated by examining replicate structure with the model Y = overall mean + replicate + error. To compare mutant and control flies, the model Y = overall + replicate (genotype) + error was used. To exclude locomotor defects that could bias the TPB assay results, we verified the motor activity of all the temperature preference flies assayed.

25. Eisel, U. *et al.* Tetanus toxin light chain expression in Sertoli cells of transgenic mice causes alterations of the *actin* cytoskeleton and disrupts spermatogenesis. *EMBO J.* **12**, 3365–3372 (1993).
26. McGuire, S. E., Roman, G. & Davis, R. L. Gene expression systems in *Drosophila*: a synthesis of time and space. *Trends Genet.* **20**, 384–391 (2004).
27. Riemensperger, T., Voller, T., Stock, P., Buchner, E. & Fiala, A. Punishment prediction by dopaminergic neurons in *Drosophila*. *Curr. Biol.* **15**, 1953–1960 (2005).

LETTERS

Essential roles of PI(3)K–p110 β in cell growth, metabolism and tumorigenesis

Shidong Jia^{1,3*}, Zhenning Liu^{1,3*}, Sen Zhang^{1,3*}, Pixu Liu^{1,3*}, Lei Zhang^{1,3}, Sang Hyun Lee^{1,3}, Jing Zhang^{1,3}, Sabina Signoretti^{2,4}, Massimo Loda^{2,4}, Thomas M. Roberts^{1,3} & Jean J. Zhao^{1,3,5}

On activation by receptors, the ubiquitously expressed class IA isoforms (p110 α and p110 β) of phosphatidylinositol-3-OH kinase (PI(3)K) generate lipid second messengers, which initiate multiple signal transduction cascades^{1–5}. Recent studies have demonstrated specific functions for p110 α in growth factor and insulin signalling^{6–8}. To probe for distinct functions of p110 β , we constructed conditional knockout mice. Here we show that ablation of p110 β in the livers of the resulting mice leads to impaired insulin sensitivity and glucose homeostasis, while having little effect on phosphorylation of Akt, suggesting the involvement of a kinase-independent role of p110 β in insulin metabolic action. Using established mouse embryonic fibroblasts, we found that removal of p110 β also had little effect on Akt phosphorylation in response to stimulation by insulin and epidermal growth factor, but resulted in retarded cell proliferation. Reconstitution of p110 β -null cells with a wild-type or kinase-dead allele of p110 β demonstrated that p110 β possesses kinase-independent functions in regulating cell proliferation and trafficking. However, the kinase activity of p110 β was required for G-protein-coupled receptor signalling triggered by lysophosphatidic acid and had a function in oncogenic transformation. Most strikingly, in an animal model of prostate tumour formation induced by *Pten* loss, ablation of p110 β (also known as *Pik3cb*), but not that of p110 α (also known as *Pik3ca*), impeded tumorigenesis with a concomitant diminution of Akt phosphorylation. Taken together, our findings demonstrate both kinase-dependent and kinase-independent functions for p110 β , and strongly indicate the kinase-dependent functions of p110 β as a promising target in cancer therapy.

Class IA PI(3)Ks are heterodimeric lipid kinases consisting of a p110 catalytic subunit complexed to one of several regulatory subunits, collectively called p85 (refs 4, 5). In response to stimulation by growth factor, p110 subunits catalyse the production of the lipid second messenger phosphatidylinositol-3,4,5-trisphosphate (PtdIns(3,4,5)P₃) at the membrane^{1–4}. This second messenger in turn activates the serine/threonine kinase Akt and other downstream effectors^{9,10}. Knockout mice for either p110 α or p110 β die early in embryonic development^{11,12}. However, recent studies using conditional knockout strategies⁸ and isoform-specific small molecule inhibitors⁷ showed that p110 α is important in growth-factor signalling, whereas a kinase-inactive knock-in mouse model showed that insulin responses depended on the catalytic activity of p110 α (ref. 6).

To investigate the role(s) of p110 β in cell, tissue and organismal physiology and to examine it as a potential therapeutic target in cancer, we generated mice carrying a conditional *Pik3cb* allele (Supplementary Fig. 1). We first investigated the role of p110 β in insulin action. Because liver is the major insulin-responsive organ,

we examined the effects of p110 β loss on hepatic insulin function. To achieve liver-specific deletion of p110 β , we injected the tail veins of p110 $\beta^{\text{flox/flox}}$ mice with adenoviruses expressing β -galactosidase (Ade-LacZ) or Cre recombinase (Ade-Cre) to generate matched cohorts of control mice and mice with hepatocyte-specific deletion of p110 β . Additional cohorts of wild-type animals were subjected to Ade-Cre or Ade-LacZ, allowing us to rule out potential non-specific Cre effects (data not shown). A more than 90% decrease in p110 β protein was seen in the livers of Ade-Cre-injected mice, whereas p110 β expression remained unchanged in the livers of the control mice and muscle tissues from both groups as measured by western blotting (Supplementary Fig. 2a, b). Consistent with previous findings^{6,7} that the kinase activity of p110 β has only a minor function in insulin signalling, we saw no significant change in Akt phosphorylation in response to insulin challenge in livers lacking p110 β (Supplementary Fig. 2a). However, mice deficient in hepatic p110 β had higher levels of insulin in the blood than control animals when fasted (Fig. 1a). These animals also showed a lower tolerance of glucose and sensitivity to insulin on challenge by intraperitoneal injection of glucose or insulin (Fig. 1b, c). Mice deficient in hepatic p110 β produced more glucose than control animals did in a pyruvate challenge test (Fig. 1d). An analysis of lipogenesis showed no significant changes in serum triglycerides, fatty acids and cholesterol levels when p110 β was deleted from liver (Supplementary Fig. 3), but leptin levels were elevated compared with those in control animals, as was seen in p110 α kinase-dead knock-in animals⁶ (Supplementary Fig. 3). Of a panel of gluconeogenic genes, only that encoding phosphoenolpyruvate carboxykinase (PEPCK) was increased in p110 β -deficient livers (Supplementary Fig. 4). PEPCK promotes the production and synthesis of glucose in liver, resulting in a greater release of glucose into blood. This result therefore provides at least a partial explanation for the metabolic phenotypes observed. Although these findings indicate that p110 β might contribute to metabolic regulation through a kinase-independent mechanism, we cannot rule out the involvement of the catalytic role of p110 β in insulin responses. Our observations are in line with earlier work⁷ in which a p110 β -specific small-molecule inhibitor was used to demonstrate that acute blockage of the kinase activity of p110 β had little effect on insulin action. In addition, another study¹³ found that mice doubly heterozygous for knockout of p110 α and p110 β showed decreased sensitivity to insulin with no apparent changes in Akt phosphorylation.

To obtain cells for detailed signalling studies, mouse embryonic fibroblasts (MEFs) were isolated from floxed embryos and their wild-type littermates, as described in Supplementary Information (Supplementary Fig. 5a–c). MEFs lacking p110 β proliferated significantly more slowly than parental (p110 $\beta^{\text{flox/flox}}$) or wild-type

¹Department of Cancer Biology and ²Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, Massachusetts 02115, USA. ³Department of Pathology and ⁴Department of Pathology and ⁵Department of Surgery, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts 02115, USA.

*These authors contributed equally to this work.

($p110\beta^{+/+}$ after Cre) MEFs (Fig. 2a). To obtain a second, more easily renewed supply of knockout cells, we established immortalized $p110\beta^{\text{lox/lox}}$ and $p110\beta^{+/+}$ MEFs by means of infection with a retrovirus encoding a dominant-negative form of p53 (DNp53)¹⁴. We also generated an add-back line by introducing haemagglutinin (HA)-tagged human $p110\beta$ to DNp53-immortalized $p110\beta^{\text{lox/lox}}$ MEFs. These immortalized MEFs were then treated with Ade-Cre to yield the following MEF lines: β KO (from $p110\beta^{\text{lox/lox}}$) and β KO+ β (from the add-back). For wild-type control MEFs, designated WT, we used DNp53-immortalized $p110\beta^{\text{lox/lox}}$ cells without Cre treatment interchangeably with Cre-treated DNp53-immortalized $p110\beta^{+/+}$ MEFs, because no significant differences were ever seen between these two possible controls. Deletion of $p110\beta$ had no obvious negative effect on the phosphorylation of Akt in either primary MEFs or DNp53-immortalized MEFs in response to stimulation by insulin, epidermal growth factor (EGF) and platelet-derived growth factor (Fig. 2c and Supplementary Fig. 6a–c). However, a moderate diminution in the phosphorylation of the S6 ribosomal protein (S6RP) at Ser 235/Ser 236 was detected in these β KO cells in response to insulin or serum (Supplementary Fig. 7). Previous studies have implicated $p110\beta$ in signalling elicited by G-protein-coupled receptors (GPCRs)^{15,16}. We found consistently that both phospho-Akt and phospho-S6RP levels were decreased in response to lysophosphatidic acid (LPA) in cells lacking $p110\beta$ (Fig. 2d and Supplementary Fig. 8a).

To dissect the potential kinase-dependent and kinase-independent roles of $p110\beta$, we reconstituted β KO MEFs with a kinase-inactive allele of HA-tagged human $p110\beta$, using the previously reported K805R mutation (KR)¹⁷ to generate the β KO+KR MEF line. Though the expression of KR was lower than that of the WT add-back construct, it was expressed at a level slightly higher than endogenous $p110\beta$, and expression levels of $p110\alpha$ were unchanged (Supplementary Fig. 9a, and data not shown). Loss of lipid kinase activity in the KR cells was confirmed by lipid kinase assay^{8,18} after immunoprecipitation with anti- $p110\beta$ (Supplementary Fig. 9b). We

then examined the effect of WT or KR add-back on the altered signalling seen after the loss of $p110\beta$. The decrease in both phospho-Akt and phospho-S6RP in response to stimulation with LPA observed in β KO cells was restored by adding back WT but not the KR allele of $p110\beta$ (Fig. 2d and Supplementary Fig. 8b), suggesting a catalytic function for $p110\beta$ in LPA signalling. This seems to be unique to $p110\beta$, because loss of $p110\alpha$ has no obvious effect on LPA signalling (Fig. 2e). The lower phospho-S6RP levels in β KO cells were restored by both WT and KR add-backs in response to insulin or fetal bovine serum (FBS) (Supplementary Fig. 7, and data not shown), suggesting a scaffolding role of $p110\beta$ in signalling by insulin and growth factors. However, our MEF data do not rule out a role for $p110\beta$ in classical signalling by PI(3)K in other circumstances. For instance, when we ablated $p110\alpha$ in our earlier work, residual phosphorylation by Akt was observed in response to growth factors⁸. Because MEFs express $p110\alpha$ and $p110\beta$ and not other class I PI(3)Ks, this residual signal was presumably transduced by $p110\beta$. We also note that $p110\beta$ ablation removes the protein as a competitor for $p110\alpha$ on receptors, which may allow any decrease in signalling caused by $p110\beta$ loss to be masked or compensated for by an increased signal flux through $p110\alpha$.

To test the kinase-dependent and/or kinase-independent effects of $p110\beta$ on cell proliferation, we studied cell cycle kinetics by first synchronizing cells by serum starvation and then measuring the proportion of cells in S phase with the incorporation of bromodeoxyuridine (BrdU) after re-feeding. Whereas β KO cells had a delayed peak of BrdU incorporation, KR reconstituted cells showed a similar incorporation of BrdU to that of WT and β KO+ β cells (Fig. 2f).

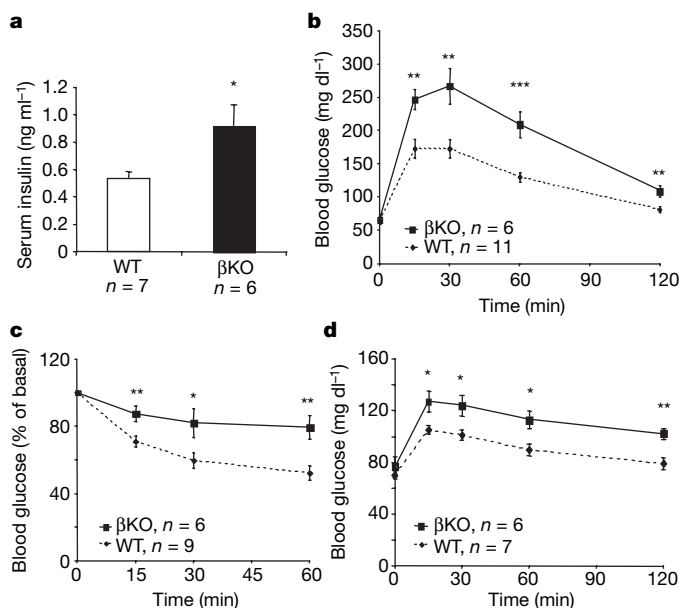


Figure 1 | Mice with liver-specific deletion of $p110\beta$ show resistance to insulin and intolerance of glucose. Mice 8–10 weeks old were injected with adenoviruses expressing LacZ or Cre recombinase. Two weeks after injection, metabolism was analysed as follows: fasted serum insulin levels (a); glucose tolerance test (b); insulin tolerance test (c) (results represent blood glucose concentrations as a percentage of starting value at time zero); pyruvate challenge (d). Data are shown as means and s.e.m. Asterisk, $P < 0.05$; two asterisks, $P < 0.01$; three asterisks, $P < 0.001$ (t -test).

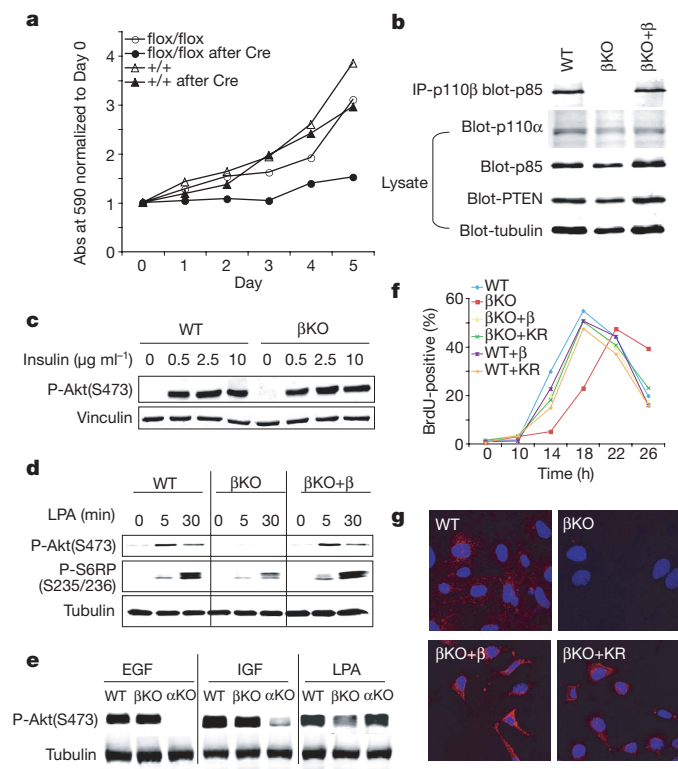


Figure 2 | Analyses of the effects of $p110\beta$ deletion on cell growth and signalling. a, Loss of $p110\beta$ retards cell growth of primary MEFs. Representative data are shown from triplicate experiments. b, Loss of endogenous $p110\beta$ protein in immortalized MEFs. IP, immunoprecipitation. c, Loss of $p110\beta$ has no negative effect on insulin signalling. d, Loss of $p110\beta$ impairs LPA-induced signalling. e, Comparison of the responses of α KO, β KO and WT MEFs to stimulation with EGF, insulin-like growth factor (IGF) or LPA. f, BrdU incorporation assay. g, Transferrin uptake in various MEF lines is shown as indicated. Transferrin is labelled red and 4,6-diamidino-2-phenylindole blue.

Consistently, β KO + KR MEFs showed proliferation rates similar to those of WT and β KO + β cells (Supplementary Fig. 9c), suggesting a kinase-independent role of p110 β in cell proliferation.

Because previous studies have found p110 β associated with members of the Rab family of small G proteins and clathrin-coated vesicles¹⁹, we measured transferrin uptake in β KO MEFs and found it to be defective compared with that in WT and β KO + β MEFs (Fig. 2g). Interestingly, normal uptake of transferrin was restored by the KR construct (Fig. 2g). Although there is ample published evidence indicating the importance of transferrin uptake for the growth of a variety of cell types²⁰, it is not clear whether the defect in transferrin uptake is a primary cause of the growth defect observed here.

Class IA PI(3)Ks have been clearly implicated in cancer^{21–24}, with much recent work delineating the role of p110 α in cancer^{25–27}. To study a potential role of p110 β in oncogenic transformation, we performed focus formation assays by infecting monolayers of DNP53-immortalized MEFs with retroviruses expressing various oncogenes. Oncogenic HRas-G12V and EGFR-Del (Δ L747–E749, A750P) retroviruses efficiently raised foci in WT cells but failed to transform β KO MEFs (Fig. 3a). The decreases in foci seen in β KO MEFs were actually more pronounced than those seen in p110 α KO MEFs (Supplementary Fig. 10). Transformation was fully restored in

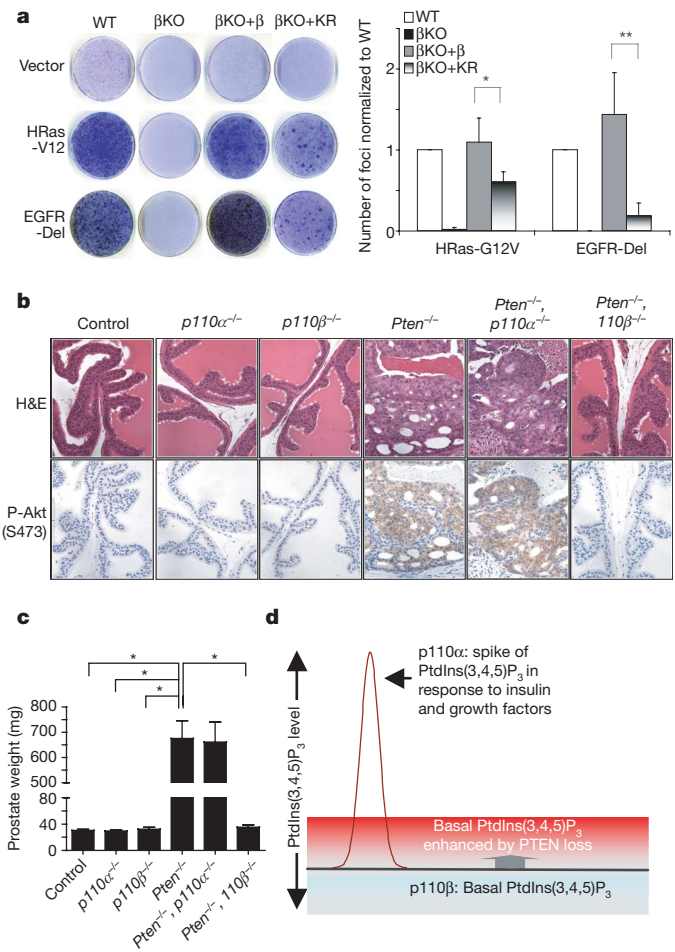


Figure 3 | Kinase activity of p110 β contributes to transformation both *in vitro* and *in vivo*. **a**, Focus formation assay in KO and reconstituted MEFs. Results are shown as means and s.e.m. for four independent experiments (asterisk, $P < 0.05$; two asterisks, $P < 0.01$, t -test). **b**, Effects of p110 β or p110 α ablation on tumorigenesis caused by PTEN loss in the anterior prostate. **c**, Quantification of the weight (means and s.e.m.) of anterior prostate tissues of the indicated strain ($n = 10$ per group; asterisk, $P < 0.001$, t -test). **d**, A model for the elevation of basal PtdIns(3,4,5)P₃ signals derived from p110 β catalytic activity induced by PTEN loss.

β KO + β cells but partly restored in β KO + KR cells (Fig. 3a), suggesting that both the kinase activity and kinase-independent functions of p110 β may have a function in oncogene-induced transformation.

PTEN, a lipid phosphatase, functions to oppose class IA PI(3)K kinase activity. Loss of *Pten* expression is a common event in many solid tumours²⁸. The key challenge is to identify which p110 isoform's catalytic activity is unshackled by *Pten* loss in any given tumour. To test for a role of p110 β in tumorigenesis driven by *Pten* loss, we generated mice that carried the *Pten*^{flax/flax} (ref. 29) and p110 β ^{flax/flax} alleles, as well as a probasin-driven Cre transgene³⁰, to specifically delete *Pten* and p110 β in prostatic epithelium. Prostates had a normal appearance in the absence of p110 β (Fig. 3b, c). Prostate tissue lacking *Pten* expression showed universal high-grade prostatic intraepithelial neoplasia in the anterior lobe by 12 weeks. Ablation of p110 β blocked the tumorigenesis caused by *Pten* loss in the anterior prostate (Table 1 and Fig. 3b, c). The loss of *Pten* was confirmed by genomic DNA analysis after laser-capture-assisted microdissection of single epithelial layers and by western blotting (Supplementary Fig. 11, and data not shown). Whereas Cre-mediated loss of *Pten* efficiently activated Akt in the prostate as judged by its phosphorylation on Ser473, additional ablation of p110 β diminished the phospho-Akt levels (Fig. 3b), suggesting that p110 β catalytic activity contributes to tumorigenesis. More surprisingly, when we performed the same set of experiments using p110 α ablation, we saw no changes either in tumour formation or in Akt phosphorylation (Table 1 and Fig. 3b, c). Again, the complete excision of p110 α in tumour tissues was confirmed by multiple measures (Supplementary Fig. 12, and data not shown). It has been suggested that p110 α and p110 β generate distinct pools of PtdIns(3,4,5)P₃ (ref. 7). In response to insulin or other stimuli, an acute flux of PtdIns(3,4,5)P₃ is produced largely by p110 α and is efficiently coupled to Akt phosphorylation. In contrast, p110 β has been proposed to generate a basal level of PtdIns(3,4,5)P₃ with little effect on Akt phosphorylation⁷. It was shown that Akt phosphorylation induced by *Pten* loss *in vitro* was sensitive to p110 β -specific inhibitors⁷. We propose that it is this basal PtdIns(3,4,5)P₃ signal that has been enhanced to drive transformation and Akt activation by *Pten* loss in the murine prostate (Fig. 3d). Alternatively, the differential effects of p110 α and p110 β ablation may arise because the signal activating PI(3)K is generated by an as yet unidentified GPCR or a receptor tyrosine kinase that functions through p110 β .

Thus, our data suggest distinct functions for p110 β and p110 α in cell signalling and transformation. We have showed that p110 β has an important physiological function in metabolic regulation and glucose homeostasis, perhaps involving a kinase-independent mechanism. A kinase-independent function of p110 β was further suggested in controlling cell proliferation and trafficking in p110 β KO MEFs and MEFs reconstituted with a WT or kinase-dead allele of p110 β . It would clearly be a mistake to overlook the contributions of p110 β as a kinase. The basal PtdIns(3,4,5)P₃ pool catalysed by p110 β seems to be 'silent' in response to stimulation by insulin and other growth factors, but becomes a 'powerhouse' to drive oncogenic

Table 1 | Effects of ablation of p110 α or p110 β on prostate tumorigenesis induced by PTEN loss

Abbreviation	Full name	Animals positive for PIN/ number of animals used
Control	Floxed littermates	0/20
p110 α ^{-/-}	p110 α ^{flax/flax} , PbCre4	0/20
p110 β ^{-/-}	p110 β ^{flax/flax} , PbCre4	0/14
Pten ^{-/-}	Pten ^{flax/flax} , PbCre4	20/20
Pten ^{-/-} ; p110 α ^{-/-}	Pten ^{flax/flax} , p110 α ^{flax/flax} , PbCre4	15/15
Pten ^{-/-} ; p110 β ^{-/-}	Pten ^{flax/flax} , p110 β ^{flax/flax} , PbCre4	0/16

Paraffin sections of anterior prostates from the indicated strains at 12 weeks were stained with haematoxylin/eosin. The pathological phenotype was uniformly observed within each genotype and is summarized in the table. PIN, prostatic intraepithelial neoplasia.

transformation in the absence of PTEN, as evident in our mouse prostate tumour model. Taken together, our findings indicate that p110 β may be an attractive target for kinase inhibitors in cancer treatment with minor metabolic disturbances.

METHODS SUMMARY

Mice carrying floxed *p110 β* (generated in this work), floxed *p110 α* (ref. 8), floxed *Pten*²⁹ and probasin-driven Cre transgene³⁰ (Mouse Models of Human Cancers Consortium (MMHCC), National Cancer Institute) were used in this study. All animals were housed and treated in accordance with protocols approved by the Institutional Animal Care and Use Committees of Dana-Farber Cancer Institute and Harvard Medical School. The generation, culture and immortalization of MEFs, growth factor signalling study, retroviral infection, cell growth, cell cycle, lipid kinase assay, transferrin internalization, focus formation, glucose tolerance testing, insulin tolerance tests, pyruvate challenge, immunoprecipitation, immunoblotting, immunohistochemical and histological analyses were performed in accordance with standard or published protocols. Statistical analyses were performed with Student's *t*-test unless otherwise indicated.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 27 February; accepted 15 May 2008.

Published online 25 June 2008.

- Vanhaesebroeck, B. & Waterfield, M. D. Signaling by distinct classes of phosphoinositide 3-kinases. *Exp. Cell Res.* **253**, 239–254 (1999).
- Blume-Jensen, P. & Hunter, T. Oncogenic kinase signalling. *Nature* **411**, 355–365 (2001).
- Vivanco, I. & Sawyers, C. L. The phosphatidylinositol 3-kinase AKT pathway in human cancer. *Nature Rev. Cancer* **2**, 489–501 (2002).
- Engelman, J. A., Luo, J. & Cantley, L. C. The evolution of phosphatidylinositol 3-kinases as regulators of growth and metabolism. *Nature Rev. Genet.* **7**, 606–619 (2006).
- Liu, Z. & Roberts, T. M. Human tumor mutants in the p110 α subunit of PI3K. *Cell Cycle* **5**, 675–677 (2006).
- Foukas, L. C. *et al.* Critical role for the p110 α phosphoinositide-3-OH kinase in growth and metabolic regulation. *Nature* **441**, 366–370 (2006).
- Knight, Z. A. *et al.* A pharmacological map of the PI3-K family defines a role for p110 α in insulin signaling. *Cell* **125**, 733–747 (2006).
- Zhao, J. J. *et al.* The p110 α isoform of PI3K is essential for proper growth factor signaling and oncogenic transformation. *Proc. Natl Acad. Sci. USA* **103**, 16296–16300 (2006).
- Bader, A. G., Kang, S., Zhao, L. & Vogt, P. K. Oncogenic PI3K deregulates transcription and translation. *Nature Rev. Cancer* **5**, 921–929 (2005).
- Vanhaesebroeck, B. *et al.* Synthesis and function of 3-phosphorylated inositol lipids. *Annu. Rev. Biochem.* **70**, 535–602 (2001).
- Bi, L., Okabe, I., Bernard, D. J., Wynshaw-Boris, A. & Nussbaum, R. L. Proliferative defect and embryonic lethality in mice homozygous for a deletion in the p110 α subunit of phosphoinositide 3-kinase. *J. Biol. Chem.* **274**, 10963–10968 (1999).
- Bi, L., Okabe, I., Bernard, D. J. & Nussbaum, R. L. Early embryonic lethality in mice deficient in the p110 β catalytic subunit of PI 3-kinase. *Mamm. Genome* **13**, 169–172 (2002).
- Brachmann, S. M., Ueki, K., Engelman, J. A., Kahn, R. C. & Cantley, L. C. Phosphoinositide 3-kinase catalytic subunit deletion and regulatory subunit deletion have opposite effects on insulin sensitivity in mice. *Mol. Cell. Biol.* **25**, 1596–1607 (2005).
- Shaulian, E., Zauberman, A., Ginsberg, D. & Oren, M. Identification of a minimal transforming domain of p53: negative dominance through abrogation of sequence-specific DNA binding. *Mol. Cell. Biol.* **12**, 5581–5592 (1992).
- Hazeki, O. *et al.* Activation of PI 3-kinase by G protein $\beta\gamma$ subunits. *Life Sci.* **62**, 1555–1559 (1998).
- Roche, S., Downward, J., Raynal, P. & Courtneidge, S. A. A function for phosphatidylinositol 3-kinase β (p85 α -p110 β) in fibroblasts during mitogenesis: requirement for insulin- and lysophosphatidic acid-mediated signal transduction. *Mol. Cell. Biol.* **18**, 7119–7129 (1998).
- Yart, A. *et al.* A function for phosphoinositide 3-kinase β lipid products in coupling $\beta\gamma$ to Ras activation in response to lysophosphatidic acid. *J. Biol. Chem.* **277**, 21167–21178 (2002).
- Nobukuni, T. *et al.* Amino acids mediate mTOR/raptor signaling through activation of class 3 phosphatidylinositol 3OH-kinase. *Proc. Natl Acad. Sci. USA* **102**, 14238–14243 (2005).
- Shin, H. W. *et al.* An enzymatic cascade of Rab5 effectors regulates phosphoinositide turnover in the endocytic pathway. *J. Cell Biol.* **170**, 607–618 (2005).
- Daniels, T. R., Delgado, T., Rodriguez, J. A., Helguera, G. & Penichet, M. L. The transferrin receptor part I: Biology and targeting with cytotoxic antibodies for the treatment of cancer. *Clin. Immunol.* **121**, 144–158 (2006).
- Bellacosa, A. *et al.* Molecular alterations of the AKT2 oncogene in ovarian and breast carcinomas. *Int. J. Cancer* **64**, 280–285 (1995).
- Li, J. *et al.* PTEN, a putative protein tyrosine phosphatase gene mutated in human brain, breast, and prostate cancer. *Science* **275**, 1943–1947 (1997).
- Steck, P. A. *et al.* Identification of a candidate tumour suppressor gene, MMAC1, at chromosome 10q23.3 that is mutated in multiple advanced cancers. *Nature Genet.* **15**, 356–362 (1997).
- Ringel, M. D. *et al.* Overexpression and overactivation of Akt in thyroid carcinoma. *Cancer Res.* **61**, 6105–6111 (2001).
- Samuels, Y. *et al.* High frequency of mutations of the PIK3CA gene in human cancers. *Science* **304**, 554 (2004).
- Zhao, J. J. *et al.* The oncogenic properties of mutant p110 α and p110 β phosphatidylinositol 3-kinases in human mammary epithelial cells. *Proc. Natl Acad. Sci. USA* **102**, 18443–18448 (2005).
- Bader, A. G., Kang, S. & Vogt, P. K. Cancer-specific mutations in PIK3CA are oncogenic *in vivo*. *Proc. Natl Acad. Sci. USA* **103**, 1475–1479 (2006).
- Chen, Z. *et al.* Crucial role of p53-dependent cellular senescence in suppression of Pten-deficient tumorigenesis. *Nature* **436**, 725–730 (2005).
- Lesche, R. *et al.* Cre/loxP-mediated inactivation of the murine Pten tumor suppressor gene. *Genesis* **32**, 148–149 (2002).
- Wu, X. *et al.* Generation of a prostate epithelial cell-specific Cre transgenic mouse model for tissue-specific gene ablation. *Mech. Dev.* **101**, 61–69 (2001).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank C. D. Stiles and J. D. Iglehart for advice, and H. Wu for providing floxed PTEN mice. This work was supported by grants from the National Institutes of Health (M.L., T.M.R. and J.J.Z.), the Department of Defense for Cancer Research (J.J.Z.), the V Foundation (J.J.Z.) and the Claudia Barr Program (J.J.Z.). In compliance with Harvard Medical School guidelines, we disclose the consulting relationships: Novartis Pharmaceuticals, Inc. (M.L., T.M.R. and J.J.Z.).

Author Contributions Z.L., S.Z. and S.L. generated the floxed *p110 β* mouse. S.J. carried out mouse tumorigenesis studies. Z.L. and S.Z. performed MEF studies. P.L. performed *in vivo* metabolic studies. L.Z. performed transferrin uptake assays. J.Z. assisted in focus formation and BrdU incorporation experiments. S.S. and M.L. performed and interpreted pathological analyses of mouse prostate tumors. T.M.R. and J.J.Z. supervised the research, interpreted the data and wrote the paper. S.J., Z.L., S.Z., P.L., L.Z., S.L. and M.L. participated in the writing of the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to T.M.R. (thomas_roberts@dfci.harvard.edu) or J.J.Z. (jean_zhao@dfci.harvard.edu)

METHODS

Mice for metabolic and tumour studies. Conditional knockout mice of *p110 β* were generated using the Cre-LoxP system. In brief, the targeting construct was assembled by isolating a 7.5-kilobase genomic fragment of *Pik3cb* from 129SvEv mouse strain and inserting two LoxP sites to flank exon 2 of *Pik3cb*. The targeting construct was electroporated into embryonic stem (ES) cells of 129SvEv mouse. ES clones carrying floxed *Pik3cb* were injected into the blastocysts of C57BL/6 mice. Male chimaeras were bred to C57BL/6 females to establish germline transmission of the conditional allele. The resulting heterozygous line (*p110 β ^{flox/+}*) was intercrossed to yield a homozygous line (*p110 β ^{flox/flox}*).

For metabolic studies, 8–10-week-old male *p110 β ^{flox/flox}* littermates were tail-vein injected with 75 μ l of adenovirus CMV-lacZ and CMV-cre (titre between 10^{10} and 4×10^{10} plaque-forming units ml^{-1} ; University of Iowa Gene Transfer Vector Core). Two weeks after injection of adenoviruses, glucose tolerance test, insulin tolerance test, pyruvate challenge and *in vivo* insulin signalling were performed as described previously³¹. Blood glucose values were determined with an Accu-Chek AVIVA glucose monitor (Roche). Serum insulin and leptin (Crystal Chem Inc.), serum-free fatty acids and triacylglycerols (Wako) and serum cholesterol (Thermo) were measured by ELISA in accordance with the manufacturer's instructions.

It took two steps to generate compound mice for tumour studies. All floxed mice used were originally in the 129-C57BL/6 background and had been backcrossed once to C57BL/6. The probasin Cre mice were in the C57BL/6 background. In step 1, male PbCre4 mice were first crossed to the female *Pten^{flox/flox}*, *p110 α ^{flox/flox}* or *p110 β ^{flox/flox}* mice, and the male offspring carrying PbCre4 were then crossed to female *Pten^{flox/flox}*, *p110 α ^{flox/flox}* or *p110 β ^{flox/flox}* mice to obtain *Pten^{flox/flox}*, PbCre4 (hereafter termed '*Pten^{-/-}*'), *p110 α ^{flox/flox}*, PbCre4 (hereafter termed '*p110 α ^{-/-}*') or *p110 β ^{flox/flox}*, PbCre4 (hereafter termed '*p110 β ^{-/-}*') mice, respectively. We used cohorts of littermates lacking the probasin Cre transgene as wild-type controls (*Pten^{flox/flox}*, *p110 α ^{flox/flox}* or *p110 β ^{flox/flox}*), and no significant differences were observed between these groups.

In step 2, *Pten^{-/-}* male mice were crossed with female *p110 α ^{flox/flox}* or *p110 β ^{flox/flox}* mice, and their male offspring carrying PbCre were crossed further with *Pten^{flox/wt}*, *p110 α ^{flox/wt}* or *Pten^{flox/wt}*, *p110 β ^{flox/wt}* littermates to obtain the desired *Pten^{flox/flox}*, *p110 α ^{flox/flox}*, PbCre4 (hereafter termed '*Pten^{-/-}*'), *p110 α ^{-/-}* or *Pten^{flox/flox}*, *p110 β ^{flox/flox}*, PbCre4 (hereafter termed '*Pten^{-/-}*'), *p110 β ^{-/-}*) compound strains. The resultant *Pten^{flox/flox}*, *p110 α ^{flox/flox}*, PbCre4 or *Pten^{flox/flox}*, *p110 β ^{flox/wt}*, PbCre4 mice from the same litters displayed high-grade prostatic intraepithelial neoplasia comparable to that of *Pten^{flox/flox}*, PbCre4 mice described in step 1, and were used interchangeably as tumour controls. Cohorts of littermates lacking the probasin Cre transgene were used as wild-type controls showing no significant differences from wild-type controls in step 1.

Alternatively, in step 2 *Pten^{-/-}*, *p110 α ^{-/-}* or *Pten^{-/-}*, *p110 β ^{-/-}* males were crossed with female littermates of *Pten^{flox/flox}*, *p110 α ^{flox/flox}* or *Pten^{flox/flox}*, *p110 β ^{flox/flox}* to generate larger numbers of the desired compound strains. No differences were seen in phenotypes in compound animals generated by either version of step 2.

Primary and immortalized MEFs. MEFs were prepared from embryos derived from intercrossing *p110 β ^{flox/+}* heterozygotes at embryonic day 13.5 after fertilization. Primary WT or floxed MEFs, and DNp53 immortalized WT or floxed MEFs were treated with Ade-Cre to generate WT control and *p110 β* -null (β KO) cells. Additional control cells used in our study were floxed MEFs without Ade-Cre treatment. β KO+ β and β KO+KR lines were generated by introducing either HA-tagged WT human *p110 β* or a kinase-inactive mutant K805R into β KO MEFs and then treated with Ade-Cre. Genotyping of MEFs was done by PCR using primer sets: LLF with LLR, and SLF with LLR (Supplementary Fig. 1).

Growth factors and western blotting. Cells were starved either for 2 h or overnight followed by stimulation with insulin (2.5 $\mu\text{g ml}^{-1}$; Sigma I2767), EGF (10 ng ml^{-1} ; Sigma E9644), IGF1 (5 ng ml^{-1} , Upstate 01-208), LPA (10 μM ; Sigma L7260) or 10% FBS for various periods as indicated in the corresponding figures. Western blot assays were performed as described previously⁸ with antibodies against PTEN (no. 9552), *p110 α* (no. 4255), phospho-Akt (Ser 473 (no. 9271) or Thr 308, no. 9275), Akt (no. 9272), phospho-p70 S6 kinase (Thr 389,

no. 9205), phospho-p44/42 MAP kinase (no. 9101), phospho-S6 ribosomal protein (Ser 235/Ser 236, no. 2211) and S6 ribosomal protein (no. 2217) (Cell Signalling), p85 (Upstate 06-195), *p110 β* (Santa Cruz, sc-602), Tubulin and Vinculin (Sigma T6199 and V9131). Immunofluorescently labelled anti-mouse IgG (Rockland Immunochemicals 610-132-003) and anti-rabbit IgG (Molecular Probes) were used to visualize western blots on an Odyssey scanner. The quantification of western blots was performed with Odyssey software version 2.0.

Growth curves. MEFs were seeded in 12-well tissue culture plates and stained with crystal violet at each indicated time point. The dye was extracted with 10% acetic acid followed by plate reading at 590 nm. The values were normalized to the absorbance at day 0. Data shown are the average of at least two independent experiments.

Cell cycle analysis. Cells were synchronized by starvation in DMEM supplemented with 0.1% FBS for 48 h before being released into the cell cycle by stimulation with 10% FBS. Cells were pulse-labelled with BrdU at each indicated time point and analysed in accordance with the manufacturer's protocol (BD Biosciences).

Focus formation assays. MEFs at 40–50% confluence were infected with various retroviruses: pBabe-Vector, pBabe-HRAS-G12V or pBabe-EGFR-Del (Δ L747–E749, A750P) and then cultured for 14–21 days for WT, β KO+ β and β KO+KR cells, but for 30–40 days for β KO cells. Confluent monolayers with foci were fixed in ethanol and stained with crystal violet.

Histology and immunohistochemistry. Prostate tissues were processed and stained as described previously³². Primary antibody used in immunohistochemistry was directed against phospho-Akt (Ser 473) (no. 3787, Cell Signalling).

Lipid kinase assays. *In vitro* lipid kinase assays were performed as described previously^{8,18}. In brief, anti-*p110 β* (Santa Cruz) immunoprecipitates from freshly prepared cell lysates were subjected to an *in vitro* lipid kinase assay with phosphatidylinositol (Avanti Polar Lipids) as the substrate. The phosphorylated lipids were resolved by thin-layer chromatography, detected by autoradiography and quantified with Adobe Photoshop.

Transferrin internalization assays. Cells were seeded on 10% poly-(L-lysine)-coated coverslips and grown overnight in DMEM medium supplemented with 10% FBS. The assay was performed as described previously³³ with Alexa Fluor555-conjugated human transferrin (Invitrogen), counterstained with 4,6-diamidino-2-phenylindole (1 $\mu\text{g ml}^{-1}$; Sigma) and mounted with mounting medium (Fisher Scientific). Cells were observed with a Zeiss confocal microscope LSM510META/NLO at $\times 63$ magnification, and images were captured with Zeiss confocal microscope software 3.2.

Laser capture microdissection and DNA extraction. Laser capture microdissection was performed as described previously³⁴. Genomic DNA of microdissected prostate epithelium was extracted with phenol/chloroform before PCR analysis.

Quantitative reverse transcription PCR analysis. Liver total RNAs were extracted with the use of the RNeasy kit (Qiagen). The following gene expression assay probes (Applied Biosystems) were used for real-time RT-PCR quantification: *phosphoenolpyruvate carboxykinase 1* (PEPCK/*Pck1*, ID no. Mm00440636_m1), *glucose-6-phosphatase* (*G6Pase*, ID no. Mm00839363_m1), *fructose bisphosphatase 1* (*Fbp1*, ID no. Mm00490181_m1), *hepatic nuclear factor 4* (*Hnf4a*, ID no. Mm00433964_m1) and *glyceraldehyde-3-phosphate dehydrogenase* (*GAPDH*, catalogue no. 4352339E). Expression was normalized to mRNA for *GAPDH* and results were expressed as fold change in mRNA compared with the indicated control mice.

31. Taniguchi, C. M. *et al.* Divergent regulation of hepatic glucose and lipid metabolism by phosphoinositide 3-kinase via Akt and PKC ζ . *Cell Metab.* **3**, 343–353 (2006).
32. Graner, E. *et al.* The isopeptidase USP2a regulates the stability of fatty acid synthase in prostate cancer. *Cancer Cell* **5**, 253–261 (2004).
33. Sever, S., Damke, H. & Schmid, S. L. Dynamin:GTP controls the formation of constricted coated pits, the rate limiting step in clathrin-mediated endocytosis. *J. Cell Biol.* **150**, 1137–1148 (2000).
34. Emmert-Buck, M. R. *et al.* Laser capture microdissection. *Science* **274**, 998–1001 (1996).

LETTERS

MicroRNAs expressed by herpes simplex virus 1 during latent infection regulate viral mRNAs

Jennifer Lin Umbach¹, Martha F. Kramer², Igor Jurak², Heather W. Karnowski¹, Donald M. Coen² & Bryan R. Cullen¹

Herpesviruses are characterized by their ability to maintain life-long latent infections in their animal hosts. However, the mechanisms that allow establishment and maintenance of the latent state remain poorly understood. Herpes simplex virus 1 (HSV-1) establishes latency in neurons of sensory ganglia, where the only abundant viral gene product is a non-coding RNA, the latency associated transcript (*LAT*)^{1,2}. Here we show that *LAT* functions as a primary microRNA (miRNA) precursor that encodes four distinct miRNAs in HSV-1 infected cells. One of these miRNAs, miR-H2-3p, is transcribed in an antisense orientation to *ICP0*—a viral immediate-early transcriptional activator that is important for productive HSV-1 replication and thought to have a role in reactivation from latency³. We show that miR-H2-3p is able to reduce *ICP0* protein expression, but does not significantly affect *ICP0* messenger RNA levels. We also identified a fifth HSV-1 miRNA in latently infected trigeminal ganglia, miR-H6, which derives from a previously unknown transcript distinct from *LAT*. miR-H6 shows extended seed complementarity to the mRNA encoding a second HSV-1 transcription factor, *ICP4*, and inhibits expression of *ICP4*, which is required for expression of most HSV-1 genes during productive infection⁴. These results may explain the reported ability of *LAT* to promote latency^{5–9}. Thus, HSV-1 expresses at least two primary miRNA precursors in latently infected neurons that may facilitate the establishment and maintenance of viral latency by post-transcriptionally regulating viral gene expression.

HSV-1 *LAT* is an ~8.3 kilobase (kb) capped, polyadenylated RNA (Fig. 1a)^{1,2} that is spliced to give a ~2.0 kb stable intron and a predicted ~6.3 kb unstable exonic RNA^{10,11}. Because *LAT* is not thought to encode a protein, we investigated whether the exonic regions of

LAT might function as a primary miRNA precursor¹². To identify HSV-1 *LAT*-derived miRNAs, we constructed a *LAT* expression plasmid, pcDNA3-*LAT*, in which a heterologous promoter drives transcription of an ~10.8 kb HSV-1 genomic fragment containing the entire 8.3 kb *LAT* (Fig. 1a). We transfected this plasmid into human 293T cells and isolated total RNA. Northern analysis showed high-level expression of the stable *LAT* intron (Fig. 1b).

Small RNAs derived from this sample were used to prepare complementary DNAs for 454 sequencing¹³. This resulted in 225,439 sequence reads (Supplementary Table 1), of which at least 144,955 represented cellular miRNAs (Supplementary Table 2A). We also recovered 651 HSV-1-derived miRNAs (Supplementary Tables 1 and 3). Six HSV-1 miRNA sequences were obtained, derived from four HSV-1 miRNA precursor hairpins (Fig. 2a). The two most common HSV-1 miRNAs were miR-H2-3p (265 reads) and miR-H4-3p (266 reads); these were derived from miRNA stem-loops that also gave rise to star strands miR-H2-5p (10 reads) and miR-H4-5p (61 reads) (Fig. 2a). We also detected miR-H3 (5 reads) and miR-H5 (40 reads). For each miRNA, HSV-1 *LAT* could be folded into the expected precursor stem-loop structure. Where both the miRNA and star strand were recovered, the characteristic ~2 nucleotide 3' overhangs were observed in the duplex intermediate (Fig. 2a).

These data show that *LAT* can be processed into miRNAs in culture but do not address expression *in vivo*. We therefore isolated small RNAs from trigeminal ganglia of mice latently infected with HSV-1 and performed deep sequencing of derived cDNAs. We obtained 254,651 sequence reads (Supplementary Table 1), of which at least 204,867 represent cellular miRNAs (Supplementary Table 2B). An extra 164 sequences represented HSV-1 miRNAs

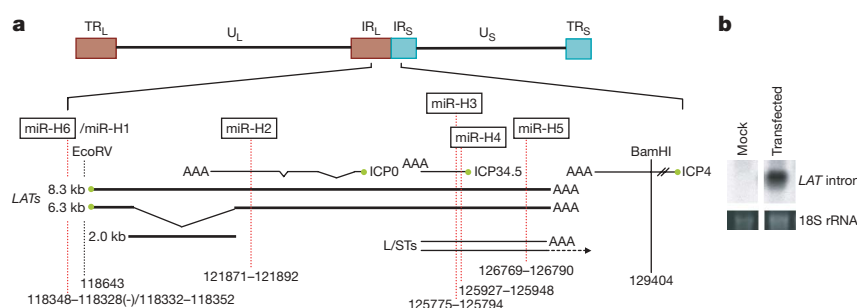


Figure 1 | Genomic location of HSV-1 miRNAs. **a**, Schematic of the HSV-1 genome expanded to show details of the *LAT* locus. Relative sizes, locations and orientations of other viral transcripts in this region are indicated. Sequence coordinates of viral miRNAs and restriction enzyme sites are given according to the HSV-1 strain 17 syn+ genome (NC_001806). All viral miRNAs are in the same orientation as *LAT* except for miR-H6. An EcoRV–BamHI fragment containing *LAT* was cloned into pcDNA3 to

generate pcDNA3-*LAT*. IR, internal repeat; TR, terminal repeat; U_L, unique long; U_S, unique short. **b**, Northern blot for the ~2.0 kb *LAT* intron, demonstrating *LAT* expression after transfection of pcDNA3-*LAT* into 293T cells. The lower bands show 18S ribosomal RNA which is a loading control. Small RNAs from this sample were used for cDNA preparation and 454 sequencing.

¹Department of Molecular Genetics and Microbiology and Center for Virology, Duke University Medical Center, Durham, North Carolina 27710, USA. ²Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, Massachusetts 02115, USA.

(Supplementary Tables 1 and 4). miR-H2-3p (94 reads), miR-H3 (18 reads) and miR-H5 (1 read) represent *LAT*-derived miRNAs previously identified in *LAT*-expressing 293T cells (Fig. 2a). However, a fourth HSV-1 miRNA, miR-H6 (50 reads), derives from an RNA stem-loop transcribed from the opposite strand of the HSV-1 genome within the *LAT* promoter (Fig. 1a). This sequence was not present in pcDNA3-LAT and therefore could not be detected in transfected 293T cells. Of the total of 171 HSV-1 short RNAs detected in trigeminal ganglia, 27 were obtained only once. Of these, 20 represent truncations or point mutants of miR-H2 to miR-H6, whereas 7 seem to represent random HSV-1 RNA breakdown products (Supplementary Table 1 and data not shown).

The identification of miR-H6 is notable for two reasons. First, miR-H6 must derive from a second HSV-1 primary miRNA precursor, distinct from *LAT*, expressed in latently-infected neurons. Although a transcript in the antisense orientation to the *LAT* promoter has been described previously¹⁴, the reported ends of this transcript exclude miR-H6. The lack of previous reports describing this primary miRNA precursor may reflect the fact that it must be cleaved to generate miR-H6 and hence is probably unstable. Second, the stem-loop that gives rise to miR-H6 lies antisense to a stem-loop transcribed from the opposite DNA strand that gives rise to a previously described HSV-1 miRNA, miR-H1 (Fig. 2a). miR-H1 is expressed late in productive replication¹⁵ and shows extensive sequence complementarity with miR-H6 (Fig. 2b). The unusual phenomenon of distinct miRNAs derived by bidirectional transcription of a single genomic locus was recently also described in mouse cytomegalovirus¹⁶.

To ascertain whether any of these HSV-1 miRNAs are expressed during productive HSV-1 infection—where *LAT* is expressed late in infection¹¹—we performed stem-loop reverse transcription followed by quantitative polymerase chain reaction (qRT-PCR) for miR-H2-3p to miR-H6 using RNA from HSV-1-infected Vero cells. The cellular miRNA let-7a was used as an internal control for RNA recovery. All five HSV-1 miRNAs were detected in infected Vero cells using qRT-PCR (Fig. 2c and Supplementary Table 5A) and/or northern analysis (Supplementary Fig. 2d). The ‘non-*LAT*’ HSV-1 miRNA miR-H6 was detected at $10^{5.0}$ molecules per ng of isolated short (<200 nucleotides) RNA, whereas the four *LAT*-derived miRNAs were detected at between $10^{2.7}$ (miR-H3) and $10^{4.1}$ (miR-H2-3p)

molecules per ng (Fig. 2c and Supplementary Table 5A). These data confirm that all five novel HSV-1 miRNAs are expressed in productively infected cells.

qRT-PCR analysis of pcDNA3-LAT-transfected 293T cells (Fig. 2c) also detected all four *LAT*-derived miRNAs, but as expected did not detect miR-H6, which is not encoded by this vector. Analysis of short RNAs derived from mouse trigeminal ganglia demonstrated the expression of all four *LAT*-derived HSV-1 miRNAs, as well as miR-H6 (Fig. 2c and Supplementary Table 5A). There is a relatively poor correlation between the levels of expression of each HSV-1 miRNA, as extrapolated from deep sequencing, when compared to the qRT-PCR analysis. This presumably reflects differences in the efficiency of cDNA synthesis.

The qRT-PCR analysis presented in Fig. 2c and Supplementary Table 5A allows us to estimate roughly how many copies of each HSV-1 miRNA are present in productively infected Vero cells versus latently infected neurons. During productive infection, miR-H1 and miR-H6 are expressed at $\sim 1,200$ and ~ 300 copies per Vero cell. In contrast, the *LAT*-derived HSV-1 miRNAs miR-H2-3p to miR-H5 are all present at <40 copies per cell (Supplementary Table 6A). These latter levels may be too low to exert a significant phenotypic effect. In latently infected trigeminal ganglia, our estimate derives from a previous report that mice latently infected with the HSV-1 strain KOS contain ~ 500 *LAT*-expressing neurons per trigeminal ganglion¹⁷. On the basis of this report, we estimate $\sim 6.3 \times 10^4$ copies per *LAT*⁺ neuron for miR-H2-3p, $\sim 4 \times 10^4$ copies per *LAT*⁺ neuron for miR-H6 and $\sim 8 \times 10^5$ copies per *LAT*⁺ neuron for miR-H4-3p. We also detected substantial amounts of miR-H4-5p ($\sim 3.2 \times 10^4$ copies per *LAT*⁺ neuron), thus suggesting that the star strand of miR-H4 might also be a functional miRNA (Fig. 2c and Supplementary Table 6C). Even if our estimate of the number of latently HSV-1 infected neurons per trigeminal ganglia is low by an order of magnitude¹⁸, the level of HSV-1 miRNAs per neuron would still be within the range of cellular miRNAs that is biologically active¹².

Although we were able to detect several different HSV-1 miRNAs in both *LAT*-expressing 293T cells and infected Vero cells, we did not detect the previously described miR-LAT¹⁹ using a range of techniques (Supplementary Figs 1 and 2). The report describing miR-LAT was recently retracted.

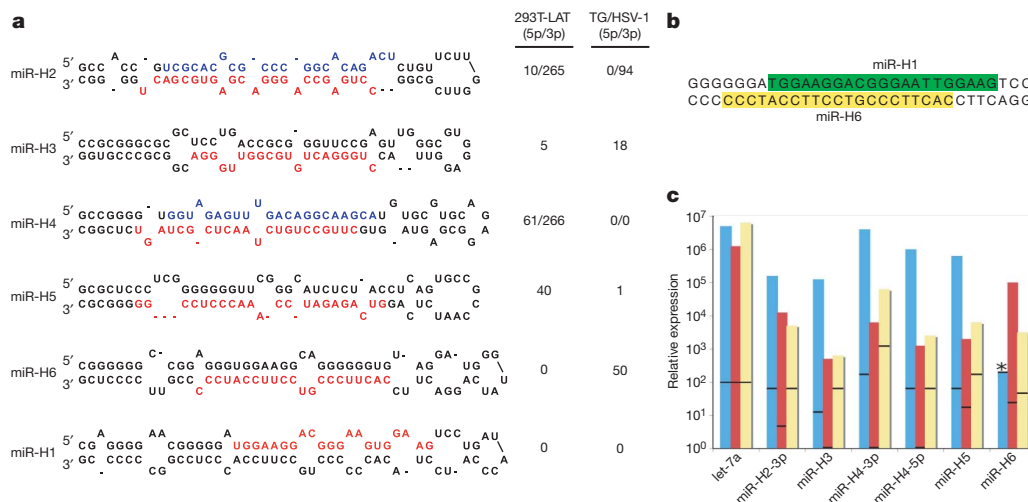


Figure 2 | HSV-1 pre-miRNAs. **a**, Predicted secondary structures of HSV-1 miRNA precursors, demonstrating the characteristic stem-loops. Mature miRNAs are indicated in red and, where observed, star strands are indicated in blue. The numbers of reads of each recovered mature miRNA sequence are indicated for *LAT*-transfected 293T cells (293T-LAT) and for trigeminal ganglia (TG) of mice infected with HSV-1. Where the star strand was also obtained, these are given as 5p/3p. miR-H1 and miR-H6 were not recovered from transfected 293T cells because pcDNA3-LAT lacks these sequences. **b**, HSV-1 genomic sequence showing the antisense orientation and overlap

of mature miR-H6, and the predicted sequence of miR-H1 (ref. 15). **c**, qRT-PCR analysis verifying the existence and relative expression of HSV-1 miRNAs in 293T cells transfected with pcDNA3-LAT (blue), Vero cells infected with HSV-1 (red), or mouse trigeminal ganglia latently infected with HSV-1 (yellow). miRNA abundances are shown as copies per ng of short-enriched RNAs (<200 nucleotides). Horizontal lines indicate background levels for each miRNA assayed. Asterisk, not detected. See Supplementary Table 5 for relevant controls.

Mapping of the six HSV-1 miRNAs onto the HSV-1 genome demonstrates that miR-H2 is antisense to the *ICP0* transcript, whereas both miR-H3 and miR-H4 are antisense to *ICP34.5* (Fig. 1a). *ICP0* is an HSV-1 transcriptional activator, expressed as an immediate-early gene, which promotes viral replication and may facilitate reactivation from latency^{3,20,21}. To examine whether miR-H2-3p could affect ICP0 protein or mRNA expression, we transfected 293T cells with either a wild-type ICP0 expression plasmid or a derivative containing three point mutations within the predicted miR-H2-3p seed region (Fig. 3a). These plasmids were co-transfected with plasmids designed to express a short hairpin RNA (shRNA) that mimics the predicted *miR-H2* pre-miRNA (Fig. 2a and Supplementary Fig. 3) or a mutated version of the *miR-H2* pre-miRNA (miR-H2-3M) that bears three mutations in the miR-H2-3p seed region that restore complementarity to the ICP0 mutant (Fig. 3a). As shown in Fig. 3b, the wild-type *miR-H2* pre-miRNA inhibited expression of wild-type, but not mutant, ICP0 protein. Conversely, expression of the mutant ICP0 protein was reduced on co-expression of the mutant *miR-H2-3M* pre-miRNA but was not affected by wild-type *miR-H2*. Although these data demonstrate that miR-H2-3p is acting through the expected target site to inhibit ICP0 protein expression, this inhibition did not correlate with a reduction in the level of *ICP0* mRNA (Fig. 3c). Similar data, obtained using short interfering RNA (siRNA) duplexes designed to mimic the miR-H2 or miR-H2-3M miRNA duplex intermediate, and using RNase protection to measure *ICP0* mRNA expression, are presented in Supplementary Fig. 3. Together, these data show that, despite the perfect complementarity of miR-H2-3p to *ICP0* mRNA, inhibition of ICP0 protein expression by this viral miRNA occurs primarily at the translational level¹². These data are consistent with earlier reports suggesting that *LAT* reduces ICP0 protein, but not mRNA, levels in infected cells^{22,23}.

Analysis of other HSV-1 genes showed sequence similarity between miR-H6, including an extended miRNA seed region¹², and the mRNA encoding ICP4, a transcription factor required for expression of most HSV-1 genes during productive infection (Fig. 4a)⁴. Co-transfection of an ICP4 expression plasmid with a synthetic form of the predicted miR-H6 duplex intermediate showed strong downregulation of ICP4 protein expression (Fig. 4b), whereas

an ICP4 expression construct with three mutations in the seed region of the predicted miR-H6 target site remained unaffected. Analysis of wild-type *ICP4* mRNA expression levels showed that miR-H6 co-expression had little or no inhibitory effect (Fig. 4c).

Here we report the identification of five HSV-1 miRNAs, three of which were previously computationally predicted^{15,24}. Four of these viral miRNAs derive from the second exon of the spliced ~6.3 kb *LAT* (Fig. 1a) and may provide both a rationale for the existence of spliced *LAT* and explain its characteristic instability^{1,11}, that is, *LAT* is probably degraded in the nucleus due to Drosha cleavage¹². In addition to the four *LAT*-derived HSV-1 miRNAs, we also identified a fifth miRNA, miR-H6, derived from an at present unknown primary miRNA precursor that lies antisense to the *LAT* promoter and which must also be expressed in latently infected neurons (Fig. 1a). Of interest, miR-H6 lies antisense to a known late HSV-1 miRNA, miR-H1 (ref. 15).

Three of the latently expressed HSV-1 miRNAs are transcribed antisense to HSV-1 mRNAs—*ICP0* mRNA in the case of miR-H2-3p and *ICP34.5* mRNA in the case of both miR-H3 and miR-H4-3p (Fig. 1a)—and we have demonstrated that miR-H2-3p is able to inhibit ICP0 protein expression (Fig. 3b). Because ICP0 is a key immediate-early HSV-1 transcriptional activator that may promote entry into the productive replication cycle^{3,20,21}, inhibition of ICP0 expression by miR-H2-3p may increase the likelihood that neurons enter and maintain latency. It has been previously proposed that *LAT* inhibits ICP0 expression post-transcriptionally in neurons^{2,10,23} and the existence of miR-H2-3p could explain this phenomenon. We also observed that miR-H6 shows partial complementarity to *ICP4* mRNA, including an extended miRNA seed region¹², and can reduce ICP4 protein expression (Fig. 4). Similar to ICP0, ICP4 can promote exit from latency²¹, and inhibition of ICP4 expression may therefore enhance the robustness of the latent state.

Although we have not directly examined the effect of miR-H3 and miR-H4-3p on *ICP34.5* expression, it seems probable that these viral miRNAs are also acting as inhibitors of viral gene expression. Data favouring this hypothesis come from analysis of the long/short junction spanning transcripts (L/STs) that overlap the 3' end of *LAT* (Fig. 1a). L/STs are expressed by HSV-1 mutants lacking ICP4 (ref. 25). Notably, the L/STs, which have the potential to give rise to miR-H3 and miR-H4-3p (Fig. 1a), are known to inhibit *ICP34.5* expression by an 'antisense' mechanism^{26,27}, and these viral miRNAs are

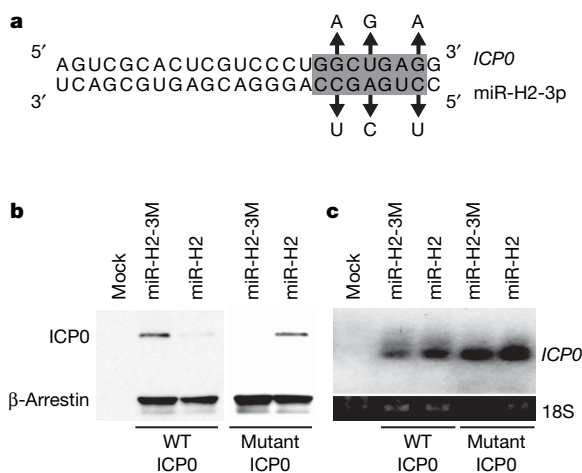


Figure 3 | Downregulation of ICP0 protein expression by HSV-1 miR-H2-3p. **a**, Sequence of miR-H2-3p bound to *ICP0* mRNA. The miRNA seed region is indicated in grey; arrows indicate complementary nucleotide changes introduced into the mutant ICP0 expression plasmid and miR-H2-3p-3M expression construct. **b**, Western blot analysis of ICP0 protein expression. 293T cells were transfected with a β -arrestin expression plasmid and either a wild-type (WT) or a mutant ICP0 expression construct. Plasmids expressing either miR-H2-3p or miR-H2-3p-3M were co-transfected. **c**, Northern analysis of the samples shown in **b**. 18S rRNA was used as a loading control.

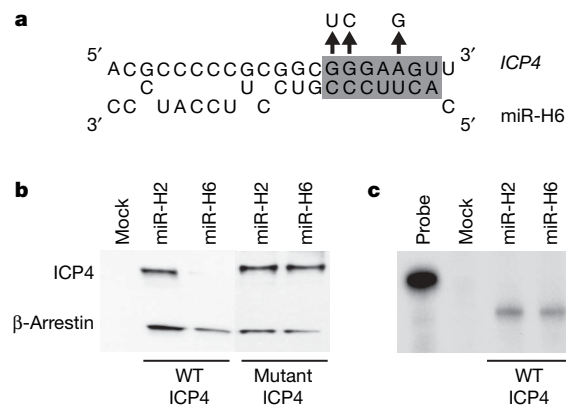


Figure 4 | Downregulation of ICP4 protein expression by HSV-1 miR-H6. **a**, Sequence complementarity of miR-H6 to nucleotides 127,298 to 127,319 of the *ICP4* mRNA. The grey box indicates the miRNA seed region; arrows indicate nucleotide changes present in the ICP4 mutant. **b**, Western blot analysis of ICP4 protein expression. 293T cells were co-transfected with a synthetic miR-H6 duplex intermediate and plasmids expressing either wild-type ICP4 or the ICP4 mutant. An miR-H2 duplex intermediate served as a negative control as *ICP4* mRNA has no predicted target sites for miR-H2-3p. **c**, RNase protection analysis of *ICP4* mRNA levels in the wild-type ICP4 samples shown in **b**.

presumably responsible for this effect. Our observation that HSV-1 miRNAs are capable of downregulating key viral immediate early proteins is consistent with the recent proposal—based primarily on computational data—that herpesviruses in general may use viral miRNAs “as part of their strategy to enter and maintain latency”²⁸.

METHODS SUMMARY

pcDNA3-LAT expresses a ~10.8 kb EcoRV to BamHI fragment, derived from the KOS strain of HSV-1, which extends 134 bp 5', and ~2.3 kb 3', to LAT. Small RNAs were prepared using standard techniques from 293T cells transfected with pcDNA3-LAT or from the dissected trigeminal ganglia of mice latently infected with HSV-1 strain KOS 30 days previously. cDNAs were prepared and subjected to 454 sequencing¹³. Vero or SY5Y cells were infected with HSV-1 strain KOS or strain 17syn+ at 10 plaque-forming units per cell, and RNA was collected for qRT-PCR analysis 14 to 18 h after infection. Northern and western blot analyses were performed using standard methods. Stem-loop qRT-PCR methods are described in Supplementary Information.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 29 April; accepted 16 May 2008.

Published online 2 July 2008.

- Bloom, D. C. HSV LAT and neuronal survival. *Int. Rev. Immunol.* **23**, 187–198 (2004).
- Stevens, J. G., Wagner, E. K., Devi-Rao, G. B., Cook, M. L. & Feldman, L. T. RNA complementary to a herpesvirus alpha gene mRNA is prominent in latently infected neurons. *Science* **235**, 1056–1059 (1987).
- Everett, R. D. ICP0, a regulator of herpes simplex virus during lytic and latent infection. *Bioessays* **22**, 761–770 (2000).
- Preston, C. M. Control of herpes simplex virus type 1 mRNA synthesis in cells infected with wild-type virus or the temperature-sensitive mutant tsK. *J. Virol.* **29**, 275–284 (1979).
- Chen, S. H., Kramer, M. F., Schaffer, P. A. & Coen, D. M. A viral function represses accumulation of transcripts from productive-cycle genes in mouse ganglia latently infected with herpes simplex virus. *J. Virol.* **71**, 5878–5884 (1997).
- Garber, D. A., Schaffer, P. A. & Knipe, D. M. A LAT-associated function reduces productive-cycle gene expression during acute infection of murine sensory neurons with herpes simplex virus type 1. *J. Virol.* **71**, 5885–5893 (1997).
- Thompson, R. L. & Sawtell, N. M. Herpes simplex virus type 1 latency-associated transcript gene promotes neuronal survival. *J. Virol.* **75**, 6660–6675 (2001).
- Thompson, R. L. & Sawtell, N. M. The herpes simplex virus type 1 latency-associated transcript gene regulates the establishment of latency. *J. Virol.* **71**, 5432–5440 (1997).
- Sawtell, N. M. & Thompson, R. L. Herpes simplex virus type 1 latency-associated transcription unit promotes anatomical site-dependent establishment and reactivation from latency. *J. Virol.* **66**, 2157–2169 (1992).
- Farrell, M. J., Dobson, A. T. & Feldman, L. T. Herpes simplex virus latency-associated transcript is a stable intron. *Proc. Natl Acad. Sci. USA* **88**, 790–794 (1991).
- Kang, W. *et al.* Characterization of a spliced exon product of herpes simplex type-1 latency-associated transcript in productively infected cells. *Virology* **356**, 106–114 (2006).
- Bartel, D. P. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* **116**, 281–297 (2004).
- Hafner, M. *et al.* Identification of microRNAs and other small regulatory RNAs using cDNA library sequencing. *Methods* **44**, 3–12 (2008).
- Perrig, G. C. *et al.* A novel herpes simplex virus type 1 transcript (AL-RNA) antisense to the 5' end of the latency-associated transcript produces a protein in infected rabbits. *J. Virol.* **76**, 8003–8010 (2002).
- Cui, C. *et al.* Prediction and identification of herpes simplex virus 1-encoded microRNAs. *J. Virol.* **80**, 5499–5508 (2006).
- Dölken, L. *et al.* Mouse cytomegalovirus microRNAs dominate the cellular small RNA profile during lytic infection and show features of posttranscriptional regulation. *J. Virol.* **81**, 13771–13782 (2007).
- Feldman, L. T. *et al.* Spontaneous molecular reactivation of herpes simplex virus type 1 latency in mice. *Proc. Natl Acad. Sci. USA* **99**, 978–983 (2002).
- Sawtell, N. M. Comprehensive quantification of herpes simplex virus latency at the single-cell level. *J. Virol.* **71**, 5423–5431 (1997).
- Gupta, A., Gartner, J. J., Sethupathy, P., Hatzigeorgiou, A. G. & Fraser, N. W. Anti-apoptotic function of a microRNA encoded by the HSV-1 latency-associated transcript. *Nature* **442**, 82–85 (2006).
- Cai, W. *et al.* The herpes simplex virus type 1 regulatory protein ICP0 enhances virus replication during acute infection and reactivation from latency. *J. Virol.* **67**, 7501–7512 (1993).
- Halford, W. P., Kemp, C. D., Isler, J. A., Davido, D. J. & Schaffer, P. A. ICP0, ICP4, or VP16 expressed from adenovirus vectors induces reactivation of latent herpes simplex virus type 1 in primary cultures of latently infected trigeminal ganglion cells. *J. Virol.* **75**, 6143–6153 (2001).
- Chen, S. H. *et al.* Neither LAT nor open reading frame P mutations increase expression of spliced or intron-containing ICP0 transcripts in mouse ganglia latently infected with herpes simplex virus. *J. Virol.* **76**, 4764–4772 (2002).
- Thompson, R. L., Shieh, M. T. & Sawtell, N. M. Analysis of herpes simplex virus ICP0 promoter function in sensory neurons during acute infection, establishment of latency, and reactivation *in vivo*. *J. Virol.* **77**, 12319–12330 (2003).
- Pfeffer, S. *et al.* Identification of virus-encoded microRNAs. *Science* **304**, 734–736 (2004).
- Yeh, L. & Schaffer, P. A. A novel class of transcripts expressed with late kinetics in the absence of ICP4 spans the junction between the long and short segments of the herpes simplex virus type 1 genome. *J. Virol.* **67**, 7373–7382 (1993).
- Randall, G. & Roizman, B. Transcription of the derepressed open reading frame P of herpes simplex virus 1 precludes the expression of the antisense γ_1 34.5 gene and may account for the attenuation of the mutant virus. *J. Virol.* **71**, 7750–7757 (1997).
- Lee, L. Y. & Schaffer, P. A. A virus with a mutation in the ICP4-binding site in the L/ST promoter of herpes simplex virus type 1, but not a virus with a mutation in open reading frame P, exhibits cell-type-specific expression of γ_1 34.5 transcripts and latency-associated transcripts. *J. Virol.* **72**, 4250–4264 (1998).
- Murphy, E., Vanicek, J., Robins, H., Shenk, T. & Levine, A. J. Suppression of immediate-early viral gene expression by herpesvirus-coded microRNAs: implications for latency. *Proc. Natl Acad. Sci. USA* **105**, 5453–5458 (2008).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank R. Sandri-Goldin for reagents used in this research and S. Boissel for contributions to PCR primer design. This work was supported by National Institutes of Health grants to B.R.C. and D.M.C.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to B.R.C. (culle002@mc.duke.edu).

METHODS

Cells, viruses and RNA. 293T and Vero cells were maintained in DMEM supplemented with 10% fetal bovine serum and 5% newborn bovine serum, respectively. SY5Y cells were maintained in RPMI supplemented with 10% fetal bovine serum. Total RNA was extracted using Trizol (Invitrogen) or mirVana miRNA Isolation Kit (Ambion). HSV-1 strain KOS was propagated and assayed as described²⁹.

Plasmid constructs, siRNAs and transfections. pcDNA3-LAT was derived from an EcoRV–BamHI digest of pSG28 (ref. 30), which released adjacent EcoRI–BamHI and BamHI–BamHI fragments. These fragments contain the entire ~8.3 kb LAT, as well as an additional ~130 bp 5' of the cap site, and ~2.3 kb 3' of the polyadenylation signal. These fragments were ligated into pcDNA3.1(–)/Zeo (Invitrogen) and screened to verify orientation. Transfection of pcDNA3-LAT into 293T cells was performed using FuGene (Roche).

The ICP0 and ICP4 expression constructs pRS-1 and pSG28K/B, respectively, have been described previously³¹. pRS-1 contains the entire ICP0 gene, whereas pSG28K/B contains the entire ICP4 gene, both including the cognate promoter and poly(A) site, inserted into pUC18.

The ICP0-3M mutant was generated by PCR using two primer pairs. The 5' half of the mutant was generated using primers 5'-TCTGTCTTGGTTCGCGGCTTGCCTGACGGACGAGTGCGACT-3' and 5'-CGCTCGAGAACAGAGACCCCATAGTGATCA-3'. The 3' half was generated using primers 5'-AGTCGCACTCGTCCCTGACGCAAGCCGCGAACCAAGAACA-3' and 5'-CGGAATCTGAGTCGGAGGGGGGTGCGTC-3'. The two PCR products were then woven together using recombinant PCR. The product was then digested with MluI/BclI and ligated into a similarly digested pRS-1 vector grown in SCS110 dam[–]/dcm[–] bacteria (Stratagene).

The ICP4 mutant was also generated by PCR. pUC18 was modified by inserting annealed oligonucleotides containing unique BsaAI/XhoI/DraIII sites into the vector HindIII/EcoRI sites. A fragment extending from the ICP4 BsaAI site to the miR-H6 binding site was then generated by PCR, which also introduced an XhoI site into the miRNA seed, using primers 5'-CCGCTGCGGC CCGGTACGTGGCGCTGGGGCGCGAGGCGGTGCGCGCCGGCCCGGCC CG-3' and 5'-CGTCCCAGGCGCCTTCCAGTCCACAACCTCGAGCCGC GGGGGCGTGGCCAAAGCCGCCT-3'. This fragment was then digested with BsaAI/XhoI and ligated into the modified pUC18. A second PCR fragment from the miR-H6 binding site to the ICP4 DraIII site was then generated using primers 5'-AGGCGGGCTTGGCCACGCCCCGCGGCTC GAGGTTGTGGACTGGGAAGGCGCCTGGGACG-3' and 5'-CGGCGCGCC AGGCGGGCGGCCGAGGCCAGACCACAGGTGGCGACCCGGACGTG GGG-3'. Like the first PCR product, this fragment also introduced an XhoI site into the miR-H6 seed. The fragment was digested with XhoI/DraIII and ligated into the pUC18 vector containing the mutant BsaAI/XhoI ICP4 fragment. The entire mutant ICP4 fragment was then transferred into pSG28K/B to generate the ICP4 mutant.

shmiR-H2-3p expression constructs were generated by annealing together oligonucleotides 5'-GATCCCCGTGCGACGCGCCTGGCATAGACTTGA CCTGAGCCAGGGACGAGTGCGACTTTT-3' and 5'-AGCTAAAAAGTCGC ACTCGTCCCTGGCTCAGGTCCAAGTCTATGCCAGGCGCGTGCACGG G-3' and ligating them into BglII/HindIII-digested pSUPER³². shmiR-H2-3p-3M was constructed similarly using oligonucleotides 5'-GATCCCCGT GCGACGCGCCTGACATAAAGTGGACTTGCCTCAGGGACGAGTGCGCA CTTTTT-3' and 5'-AGCTAAAAAGTCGCACTCGTCCCTGACGCAAGTCC AAGTTATGTGAGGGCGCGTGCACGGG-3'.

The ICP0 riboprobe used for northern blot analysis was generated by *in vitro* transcription of linearized pcDNA3.1(–)/Zeo vector containing a 431 bp BamHI–XhoI ICP0 gene fragment. The ICP0 riboprobe used for RPA was generated by *in vitro* transcription of a linearized pcDNA3 vector containing a 284 bp fragment of ICP0 that overlaps the miR-H2 binding site. The ICP4 riboprobe was generated by *in vitro* transcription of a linearized pcDNA3 vector containing a 194 bp NotI/BamHI fragment of ICP4. The HA-tagged β -arrestin expression plasmid has been described previously³³.

siRNA duplexes designed to mimic the miR-H2 and miR-H6 miRNA duplex intermediates, and a mutant form with 3 mutations (3M) in the miR-H2-3p seed region, were obtained from IDT. miR-H2-3p duplex 5' arm 5'-UCGCACUCGUCCUGGCUAGACU-3'; miR-H2-3p duplex 3' arm 5'-CCUGAGCCAGGGACGAGUGCGACU-3'; miR-H2-3p-3M duplex 5' arm 5'-UCGCACUCGUCCUGAGCGCAACU-3'; miR-H2-3p-3M duplex 3' arm 5'-CUUGCGUCAGGGACGAGUGCGACU-3'; miR-H6 duplex 5' arm 5'-GAUGGAAGGACGGGAAGUAUA-3'; miR-H6 duplex 3' arm 5'-CACUUCGGUCCUCCAUAUCC-3'. All plasmid DNA and siRNA cotransfections were performed in 293T cells using Lipofectamine 2000 (Invitrogen).

Briefly, 293T cells were plated in 24-well plates to be 80%–90% confluent on the next day. Cells were co-transfected with either 60–80 ng of an ICP0 expression plasmid or 40–60 ng of an ICP4 expression plasmid, 30 ng of p β -arrestin-HA, and 10 pmol of the appropriate siRNA duplex or 100 ng of the relevant shRNA expression vector. One microlitre of Lipofectamine 2000 was used per well per transfection. Samples were transfected in duplicate and collected simultaneously ~24 h after transfection. One sample was collected for western blot and the other for RNA analysis.

454 sequencing. Sample preparation for 454 sequencing was conducted as described³⁴, up to and including the RT–PCR step. After that point, the protocol as outlined by the Hannon laboratory on the 454 website (<http://www.454.com>) was followed. Initially, 750 μ g of total RNA from pcDNA3-LAT transfected 293T cells, and 60 μ g of total RNA from latently-infected mouse trigeminal ganglia, was used. Data analysis was performed using Microsoft Excel and BLAST. Briefly, raw data from 454 was indexed and binned to give a list and count of all unique sequences. The 5' and 3' linker sequences were parsed out and final sequences 18–24 nucleotides in length occurring twice or more were analysed by BLAST (v2.2.17) against either the mature human or the mouse miRNA databases downloaded from miRBASE (release 9.2). Sequences identified as host miRNAs were annotated, indexed and removed from subsequent analysis. Remaining sequences were then analysed against the HSV-1 genome (NC_001806) to identify potential HSV-1 miRNAs. Candidate HSV-1 miRNA sequences and flanking regions were analysed by mfold to predict the secondary structures at each locus.

Northern blots, RPA and splint-ligation assay. For the ICP0 northern blot, 5 μ g of total RNA was separated on a 0.6% agarose gel, transferred onto nitrocellulose, ultraviolet-irradiated and hybridized with the ICP0 riboprobe at 68 °C and washed at 85 °C, according to standard protocols. Bands were visualized by autoradiography. The LAT intron was detected by northern blot using an oligonucleotide probe, as described³⁵. ICP4 and ICP0 mRNA levels were measured using an Ambion HybSpeed RPA kit with 5 μ g of total RNA. Splint-ligation assays (USB) were performed using 12 μ g of total RNA per sample.

Western blots. Samples were collected and run on 4–20% Tris-HCl gels (Bio-Rad), which were then transferred onto nitrocellulose and probed with mouse monoclonal antibodies specific for ICP0 or ICP4 (Virusys), or anti-HA monoclonal antibody (Covance). Blots were then incubated with an anti-mouse secondary antibody (GE Healthcare) and bands were visualized with Lumi-Light Western Blotting Substrate (Roche).

Latent HSV infection in mice. Procedures involving mice were approved by the Harvard Medical School Institutional Animal Care and Use Committee in accordance with federal guidelines. Male CD-1 mice were infected or mock infected, housed for 30 days, and then killed for tissue as previously described³⁶.

Stem-loop qRT–PCR. Real time quantitative RT–PCR assays were designed for each HSV-1 miRNA and a cellular miRNA, let-7a, with specific stem-loop transcription primers and PCR reagents, as described³⁷. RNA standards, stem-loop RT primers and PCR primers were purchased from IDT, and TaqMan probes were purchased from Applied Biosystems (Supplementary Table 5B). Each assay was performed alongside assays of serial dilutions of each target synthetic standard miRNA, and, as a negative control, a pool of all of the other synthetic miRNAs at a concentration corresponding to the lowest dilution of the target miRNA. Furthermore, each assay included reverse transcriptase-negative and reverse transcriptase primer-negative controls. RNA was reverse transcribed in duplicate with Multiscribe (Ambion) and miRNA-specific RT primers. Aliquots of cDNA were assayed on a PRISM 7700 Sequence Detection System (Applied Biosystems). The dynamic range of each assay exceeded five orders of magnitude. To ensure quantification of low levels of mature HSV-1 miRNAs, we first isolated <200-nucleotide-long species (short-enriched) with the mirVana miRNA Isolation Kit (Ambion). RNA was fractionated using a flashPAGE System (Ambion), enriching for <40 nucleotide RNAs. Each <40 nucleotide RNA fraction was assayed for at least seven miRNAs, in duplicate. let-7a was assayed in both the short-enriched and <40 nucleotide RNAs. The number of let-7a molecules per ng of short-enriched RNA was determined (Supplementary Table 5A), and used to normalize the recovery of HSV-1 miRNAs in the <40 nucleotide RNA.

29. Coen, D. M., Fleming, H. E. Jr, Leslie, L. K. & Retondo, M. J. Sensitivity of arabinosyladenine-resistant mutants of herpes simplex virus to other antiviral drugs and mapping of drug hypersensitivity mutations to the DNA polymerase locus. *J. Virol.* **53**, 477–488 (1985).
30. Goldin, A. L., Sandri-Goldin, R. M., Levine, M. & Glorioso, J. C. Cloning of herpes simplex virus type 1 sequences representing the whole genome. *J. Virol.* **38**, 50–58 (1981).
31. Sekulovich, R. E., Leary, K. & Sandri-Goldin, R. M. The herpes simplex virus type 1 α protein ICP27 can act as a *trans*-repressor or a *trans*-activator in combination with ICP4 and ICP0. *J. Virol.* **62**, 4510–4522 (1988).

32. Brummelkamp, T. R., Bernards, R. & Agami, R. A system for stable expression of short interfering RNAs in mammalian cells. *Science* **296**, 550–553 (2002).
33. Wiegand, H. L., Doebele, B. P., Bogerd, H. P. & Cullen, B. R. A second human antiretroviral factor, APOBEC3F, is suppressed by the HIV-1 and HIV-2 Vif proteins. *EMBO J.* **23**, 2451–2458 (2004).
34. Cai, X. *et al.* Kaposi's sarcoma-associated herpesvirus expresses an array of viral microRNAs in latently infected cells. *Proc. Natl Acad. Sci. USA* **102**, 5570–5575 (2005).
35. Alvira, M. R., Goins, W. F., Cohen, J. B. & Glorioso, J. C. Genetic studies exposing the splicing events involved in herpes simplex virus type 1 latency-associated transcript production during lytic and latent infection. *J. Virol.* **73**, 3866–3876 (1999).
36. Leib, D. A. *et al.* Immediate-early regulatory gene mutants define different stages in the establishment and reactivation of herpes simplex virus latency. *J. Virol.* **63**, 759–768 (1989).
37. Chen, C. *et al.* Real-time quantification of microRNAs by stem-loop RT-PCR. *Nucleic Acids Res.* **33**, e179 (2005).

Crystal structure of the neurotrophin-3 and p75^{NTR} symmetrical complex

Yong Gong^{1*}, Peng Cao^{1*}, Hong-jun Yu¹ & Tao Jiang¹

Neurotrophins (NTs) are important regulators for the survival, differentiation and maintenance of different peripheral and central neurons. NTs bind to two distinct classes of glycosylated receptor: the p75 neurotrophin receptor (p75^{NTR}) and tyrosine kinase receptors (Trks). Whereas p75^{NTR} binds to all NTs, the Trk subtypes are specific for each NT^{1,2}. The question of whether NTs stimulate p75^{NTR} by inducing receptor homodimerization is still under debate. Here we report the 2.6-Å resolution crystal structure of neurotrophin-3 (NT-3) complexed to the ectodomain of glycosylated p75^{NTR}. In contrast to the previously reported asymmetric complex structure, which contains a dimer of nerve growth factor (NGF) bound to a single ectodomain of deglycosylated p75^{NTR} (ref. 3), we show that NT-3 forms a central homodimer around which two glycosylated p75^{NTR} molecules bind symmetrically. Symmetrical binding occurs along the NT-3 interfaces, resulting in a 2:2 ligand–receptor cluster. A comparison of the symmetrical and asymmetric structures reveals significant differences in ligand–receptor interactions and p75^{NTR} conformations. Biochemical experiments indicate that both NT-3 and NGF bind to p75^{NTR} with 2:2 stoichiometry in solution, whereas the 2:1 complexes are the result of artificial deglycosylation. We therefore propose that the symmetrical 2:2 complex reflects a native state of p75^{NTR} activation at the cell surface. These results provide a model for NTs–p75^{NTR} recognition and signal generation, as well as insights into coordination between p75^{NTR} and Trks.

NGF, brain-derived neurotrophic factor (BDNF) and NT-3 are members of neurotrophin family, which engage two types of single-transmembrane cell-surface receptor Trks and p75^{NTR} to perform a wide variety of functions in the mammalian nervous system. TrkA, TrkB and TrkC are the cognate receptors of NGF, BDNF and NT-3, respectively, and NT-3 also binds to TrkA and TrkB with low affinity^{1,2}. p75^{NTR} is a 75-kDa glycoprotein, which can bind all NTs and proneurotrophins (proNTs). p75^{NTR} belongs to the tumour necrosis factor receptor (TNFR) superfamily, which is structurally characterized by having extracellular cysteine-rich domains (CRDs) and an intracellular death domain. Signalling pathways mediated by p75^{NTR} promote either cell survival or, paradoxically, cell apoptosis. p75^{NTR} can function as a positive modulator of Trks by creating high-affinity binding sites for NTs⁴. However, the binding mode between NTs and p75^{NTR} and the crosstalk mechanism between Trk and p75^{NTR} have remained unknown. Whereas one crystallography study indicated a 2:1 asymmetric binding complex formed between NGF and the ectodomain of deglycosylated p75^{NTR} (dg-p75^{NTR})³, another biochemical report indicated that a symmetrical 2:2 stoichiometric binding complex formed between NGF and glycosylated p75^{NTR} (ref. 5). To elucidate the interaction mechanism between p75^{NTR} and NTs, we crystallized the ectodomain of glycosylated p75^{NTR} in complex with NT-3.

Crystal complexes were prepared from recombinant human NT-3 and the ectodomain of rat glycosylated p75^{NTR} (size-exclusion chromatography is shown in Supplementary Fig. 1), and the crystal structure has been determined to 2.6 Å resolution. In contrast with previous reports, we found that the NT-3–p75^{NTR} complex contains a central NT-3 homodimer with two symmetrically arranged p75^{NTR} molecules in the clefts between the NT-3 subunits (Fig. 1). The complex has a perfect non-crystallographic two-fold symmetry axis, which we suggest lies perpendicular to the cell membrane, and the root mean square deviation (r.m.s.d.) value between the two sides of the dimer is about 0.30 Å. Each p75^{NTR} protomer is about 110 Å long, and the amino and carboxy termini protrude beyond both the top and bottom ends of the NT-3 dimer. The two p75^{NTR} protomers have parallel conformations, with four kinked CRDs (CRD1–CRD4) arranged in tandem. Twelve pairs of disulphide bonds are evenly spaced along p75^{NTR}. CRD1, which includes Asn 32, is located distal to the cell membrane. CRD2, CRD3 and CRD4 interact the most with NT-3. The two p75^{NTR} C termini extend towards each other with a minor interface of about 95 Å².

A total solvent-accessible surface of about 2,314 Å² is buried between the NT-3 dimer and each copy of p75^{NTR}. A detailed analysis of the ligand–receptor contacts is shown in Fig. 2 and Supplementary Table 1. The interface can be divided into three main contact sites on p75^{NTR} that are stabilized by hydrophobic interactions, salt bridges and hydrogen bonds. The junction regions between CDR1 and CDR2, with CRD2 taking the predominant role, create site 1. Site 1 is an extensively hydrogen-bonded network, containing five hydrogen bonds and one salt bridge (Fig. 2a). Site 2 is formed by equal contributions from CDR3 and CRD4 and involves two salt bridges and two hydrogen bonds (Fig. 2b). Site 3 includes only one salt bridge between Lys 73A of NT-3 and Glu 143 of the p75^{NTR} C-terminal loop

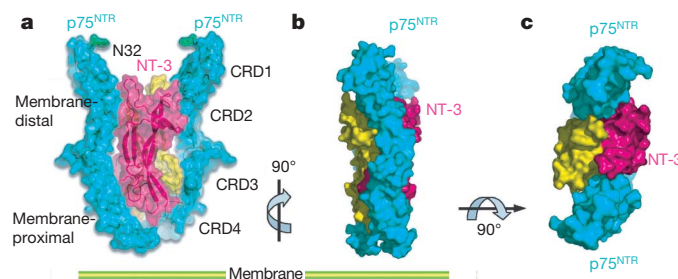


Figure 1 | Architecture of the NT-3–p75^{NTR} complex. **a**, Ribbon and surface diagram of the complex. One NT-3 monomer is shown in red and the other in yellow, and p75^{NTR} is shown in blue. The N-linked carbohydrates at Asn 32 of p75^{NTR} are shown in green. **b**, Surface representation equivalent of **a**, rotated by 90° around the vertical. **c**, As in **b**, but rotated by 90° around the horizontal.

¹National Key Laboratory of Biomacromolecules, Institute of Biophysics, Chinese Academy of Sciences, 15 Datun Road, Chaoyang District, Beijing 100101, China.

*These authors contributed equally to this work.

(Fig. 2c). Sequence alignments (Fig. 2d) show that the p75^{NTR} contact residues are, for the most part, conserved across all NTs.

In contrast to other members of the TNFR superfamily, such as TNF β –TNFR⁶ and TRAIL–DR5 (ref. 7), which are activated as pre-formed trimeric receptor complexes by trimeric ligands, p75^{NTR} is induced to form a dimer by dimeric NT-3 ligand. However, they share a common binding style in which the receptors bind along the seam of the interfaces between two ligands to form symmetrical complexes.

A detailed comparison with the structure of the unbound homodimer NT-3 (ref. 8) shows that conformational changes are induced by the binding of p75^{NTR} (Fig. 3a). First, the surface area buried at the interface of the two NT-3 monomers in the complex structure is about 2,960 Å², which is slightly smaller than the 3,190 Å² surface area in the unbound state. It results in an r.m.s.d. of 1.21 Å of the whole dimer to avoid short contacts with the receptors. Second, a prominent conformational change occurs at residues 5–12 in the N-terminal loop, which was shown to have a key function in receptor binding⁹. In the bound state, the tail of NT-3 sways to form a critical part of the interface with p75^{NTR} through hydrogen bonds and hydrophobic interactions. The other conformational differences are situated mainly at the interface with p75^{NTR}, such as movements in loops 40–48 and 70–76 and the residues Arg 31, Glu 59, Asn 65, Leu 96 and Arg 114.

Trks are known to dimerize by means of a 2:2 NT-induced clustering activation mechanism that triggers the signal transduction cascade^{10,11}. The molecular arrangements of TrkA and p75^{NTR} on NTs have opposite orientations. Although both receptors bind to the seam of the NT dimer, TrkA interacts with the middle concave face, whereas p75^{NTR} binds separately to the two convex faces (Fig. 3b–d). The buried face between the NT-3 dimer and p75^{NTR} (2,314 Å²) is very similar to that between NGF and TrkA (2,245 Å²), in line with the observation that both receptors bind all NTs with an equilibrium binding constant when expressed alone¹. The N-terminal tails of the NTs adopt quite different conformations and directions in both structures, permitting specificity of NT interaction for different receptors.

Comparison of our structure with a previous asymmetric NGF–p75^{NTR} complex³ reveals differences. First, NGF is a distorted dimer in the asymmetric complex, which leads to the hypothesis that p75^{NTR} binding to one side of NGF induces a bending of the entire NGF and disabling binding of a second p75^{NTR} to the opposite site. In contrast, NT-3 is a perfect homodimer in our structure, and the p75^{NTR} receptors bind symmetrically to form a 2:2 cluster. Second, we observed significant differences in conformation between the p75^{NTR} molecules in two structures (Fig. 4a). The most prominent conformational change concerns the N-terminal CRD1 and CRD2 of

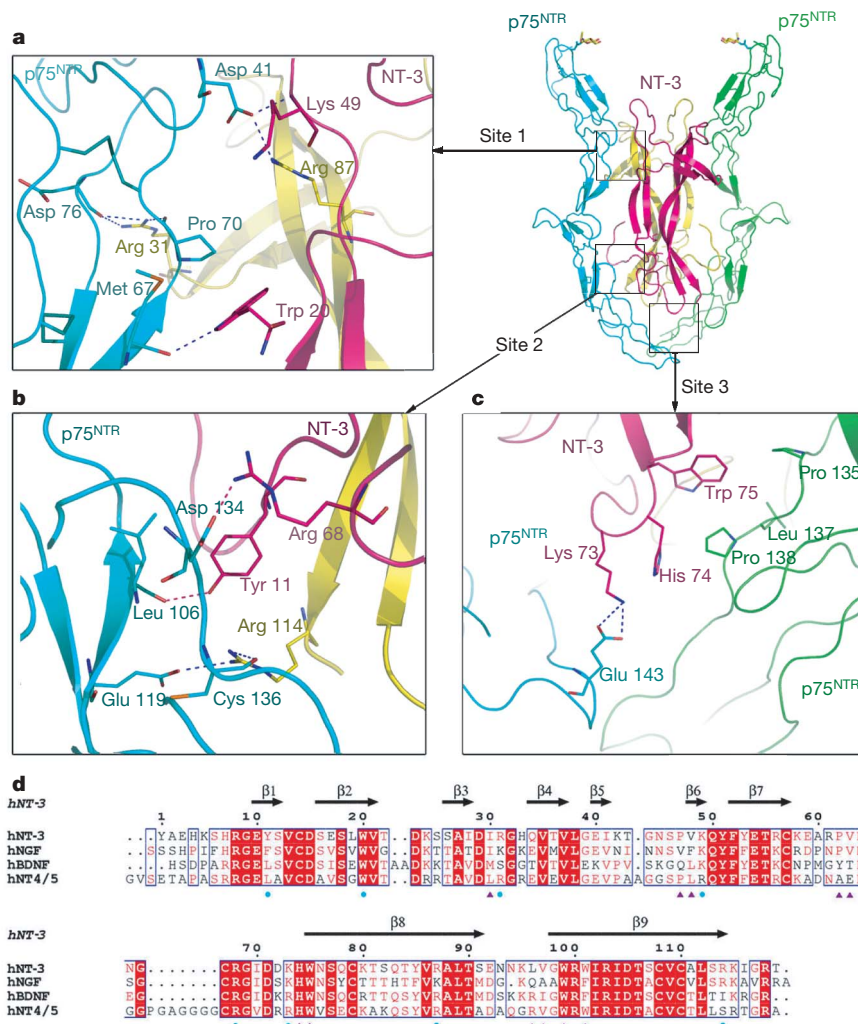


Figure 2 | Interactions between p75^{NTR} and NT-3. **a**, Close-up of the site 1 interface. NT-3 monomers are shown in red and yellow, and p75^{NTR} is shown in blue and green. Hydrogen bonds and salt bridges are shown as dashed lines. **b**, **c**, Close-ups of the site 2 (**b**) and site 3 (**c**) interfaces. **d**, Human (h) neurotrophin family sequence alignments for NT-3, NGF,

BDNF and NT4/5. Strictly conserved and conservatively substituted residues are boxed and indicated by a red background or red letters, respectively. Hydrogen bonds and salt bridges with p75^{NTR} are marked below the alignments by blue dots, and hydrophobic contacts are indicated by purple triangles.

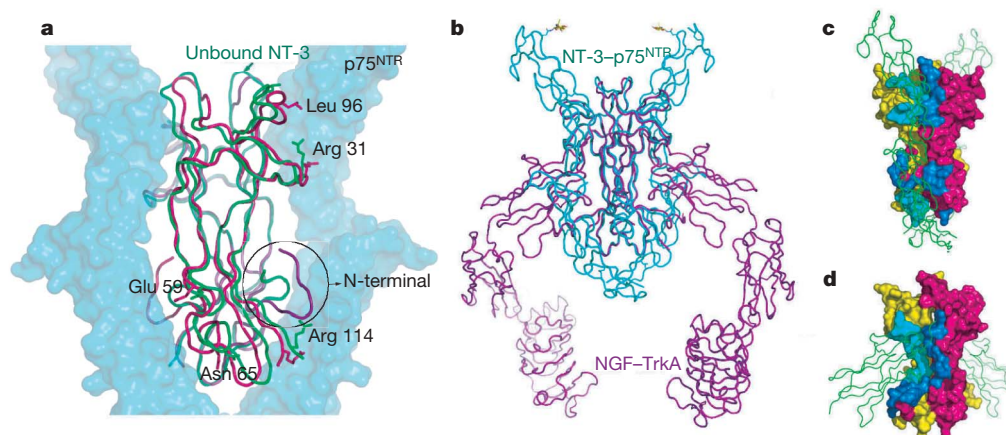


Figure 3 | Structural comparison. **a**, Comparison of bound and unbound human NT-3. Unbound human NT-3 is shown in green, bound NT-3 in red and purple, and p75^{NTR} in blue. The N terminus is circled, and the labelled amino acids show notably different conformations. **b**, Comparison of the NT-3–p75^{NTR} complex (blue) with the NGF–TrkA ectodomain complex

(purple). **c**, **d**, The regions at the surfaces of NT-3 (**c**) and NGF (**d**) that are buried by p75^{NTR} and TrkA, respectively. NT monomers are shown in red and yellow, and p75^{NTR} or TrkA is shown in green. The light and deep blue areas indicate the binding surfaces on each monomer.

p75^{NTR} with an r.m.s.d. of about 3.5 Å (Fig. 4b), resulting in changes in the interacting pairs of residues that bind NTs. Despite a better agreement in CRD3 and CRD4 (r.m.s.d. about 0.7 Å; Fig. 4c), there are also structural differences in the C-terminal loops (residues 140–152). Third, our structure is tethered together through three separate binding sites on p75^{NTR}, whereas the asymmetric complex is tethered through only two sites (sites 1 and 2). At site 1, our structure has two more hydrogen bonds and one less salt bridge than the asymmetric structure. Site 2 has an additional hydrogen bond compared with the asymmetric structure between Leu 106 of p75^{NTR} and Tyr 11B of NT-3. This hydrogen bond is determined by NT-3 because the corresponding residues to Tyr 11 are Phe 12, Leu 10 and Leu 14 on NGF, BDNF and NT4/5, respectively. The largest difference in two structures is at site 3. The C-terminal loop of p75^{NTR} in our structure sways close to NT-3, allowing the formation of a salt bridge with the 'Lys 73–His–Trp 75' loop in NT-3, which is conserved in NTs. No interaction closer than 6 Å was observed at this binding site in the asymmetric structure.

Previous studies indicated that p75^{NTR} binds to NGF and NT-3 with similar affinities¹²; however, it was still unclear whether differences between NGF and NT-3 could alter the binding stoichiometry. We measured the ligand–receptor stoichiometry of p75^{NTR} in a complex with NGF or NT-3 in solution by analytical ultracentrifugation. The results indicate that both NGF and NT-3 bind to p75^{NTR} with a 2:2 stoichiometry (Fig. 4d and Supplementary Fig. 2). These results are consistent with a report⁵ showing a 2:2 stoichiometry of human p75^{NTR} and NGF with the use of mass spectrometry, analytical ultracentrifugation and solution X-ray scattering measurements.

Glycans have a pivotal function in protein folding, oligomerization, sorting and transport in the endoplasmic reticulum and in the early secretory pathway¹³. In addition, the importance of p75^{NTR} N-linked glycosylation in NGF signalling has been demonstrated in PC12 cells¹⁴. We therefore investigated the influence of glycosylation on p75^{NTR} by expressing dg-p75^{NTR} and using Biacore analysis to measure its association constant with NT-3. We found that dg-p75^{NTR} had a significantly lower affinity than p75^{NTR} for NT-3

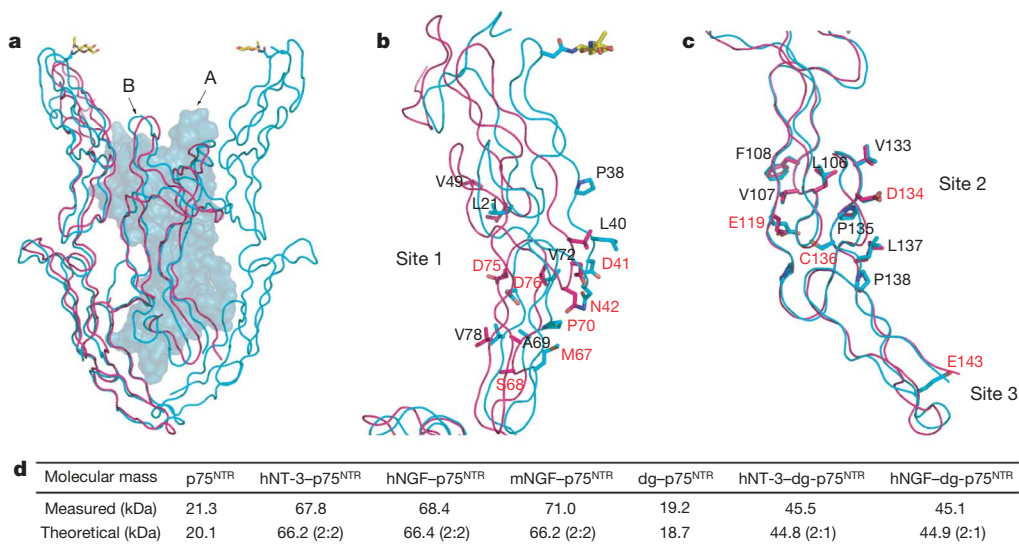


Figure 4 | Comparison of the 2:2 NT-3–p75^{NTR} complex with the 2:1 NGF–dg-p75^{NTR} complex. **a**, An overall view of the 2:2 complex (blue) and the 2:1 complex (red). Monomer A of NT-3 is shown as a molecular surface. **b**, **c**, A detailed comparison of the p75^{NTR} conformation and interactions with NT-3 or NGF. **b**, The N-terminal half of p75^{NTR}. **c**, The C-terminal half

of p75^{NTR}. The residues involved with hydrophobic interactions and hydrogen bonds (or salt bridges) are labelled in black and red, respectively. **d**, Analytical ultracentrifugation analysis of p75^{NTR}, dg-p75^{NTR} and complexes with NTs; h, human; m, mouse. The measured and theoretical molecular masses are indicated.

(Supplementary Fig. 3). Using analytical ultracentrifugation, we detected unliganded dg-p75^{NTR} as a mixture of monomers and dimers (Supplementary Fig. 4), in contrast to unliganded p75^{NTR}, which we detected as strictly monomers (Supplementary Fig. 2). These data suggest that dg-p75^{NTR} has different properties from those of native p75^{NTR}.

The stoichiometries of dg-p75^{NTR} in complex with NT-3 or NGF were further examined by analytical ultracentrifugation. For both NT-3–dg-p75^{NTR} and NGF–dg-p75^{NTR} we observed only the 2:1 complexes (Fig. 4d and Supplementary Fig. 4). These results are consistent with a previous report³ and indicate that formation of the 2:1 complex is related to receptor deglycosylation. We therefore conclude that the 2:2 ligand–receptor stoichiometry is the native and common mode of both NT-3 and NGF binding to glycosylated p75^{NTR}, whereas the 2:1 complexes are the result of artificial deglycosylation.

In our structure, only the first sugar (GlcNAc) of the five to eight carbohydrates in the N-glycan of Asn 32 was well defined by the electron density map. Although GlcNAc was not observed to make direct contacts with NT-3, the entire sugar chain probably has a function in the structural conformation and stabilization of p75^{NTR}, and in turn can affect its ligand binding activity. Similar cases were also reported for the structures of TrkA complexed with NGF¹¹ and the natriuretic peptide receptor with C-type natriuretic peptide¹⁵, in which the carbohydrate moieties on the receptors do not contact the ligands but modulate the ligand-binding activities of receptors.

The 2:2 architecture of the complex establishes a NT-induced dimerization model for p75^{NTR} activation. Two receptors bind along the interface surface grooves symmetrically, resulting in a 2:2 ligand–receptor cluster. The role of NT-3 in receptor dimerization seems to be critical, because there is little protein–protein contact between two p75^{NTR} receptors in our structure but significant contacts occur between NT-3 and p75^{NTR}, which facilitates the homodimerization of p75^{NTR}. The 2:2 cluster may result in bringing the C termini of two receptors into such close contact that the cytoplasmic regions, particularly the death domains, come into contact. This process facilitates the recruitment of intracellular interactors to generate signals downstream. Receptor clustering has been recognized as a general signal transduction mechanism for growth-factor receptors¹⁶. The high structural conservation of NTs (Fig. 2d) and biochemical results lead us to propose that the 2:2 symmetrical complex reflects the native activated state of p75^{NTR} at the cell surface when bound to NTs, providing a ubiquitous model for p75^{NTR} recognition, activation and signal generation.

The crosstalk between p75^{NTR} and Trks creates high-affinity NT-binding sites on neuronal cells and alters the signalling properties of both partners through poorly understood mechanisms¹⁷. However, the superposition of our structure on the crystal structure of the ectodomain of TrkA and NGF complex^{10,11,18} shows that there are mutually exclusive binding sites on the ligands. It is therefore impossible for a p75^{NTR}–NT–TrkA ternary complex to form a high-affinity binding site or an instantaneous intermediate binding state through interactions between extracellular domains of the receptors. Instead, the ectodomains of NT receptors simply share a ligand-competing relationship, rather than direct interactions. The high-affinity binding site probably occurs through the cytosolic and transmembrane domains of p75^{NTR} and Trks. These interactions may be not direct but facilitated by the pool of intracellular interaction partners and transmembrane constituents, which form a bridge between p75^{NTR} and Trks.

Although the precise mechanism of how limited combinations of distinct NTs and their receptors can produce such a multitude of specific cell responses is still unknown, we propose that the current 2:2 symmetrical NT-3–p75^{NTR} complex represents a native active state for NT-3–p75^{NTR} interaction during neuronal development. This symmetrical NT-3–p75^{NTR} complex provides a molecular basis for the recognition of receptors for NTs and signal transduction

mechanisms. Our structure also provides critical insights, which may allow the design of specific NTs agonists for the fine tuning of neuronal development.

METHODS SUMMARY

Protein expression. The rat p75^{NTR} ectodomain, which shares 94% identity with human p75^{NTR}, was expressed in sf9 cells. The secreted soluble p75^{NTR} protein was N-glycosylated at residue Asn 32 of the mature protein. The protein was purified by affinity chromatography and gel filtration. Recombinant human NT-3 and human NGF were expressed in *Escherichia coli* (Genentech). To form the complex, p75^{NTR} (20 mg ml^{−1}) and NT-3 (20 mg ml^{−1}) were mixed at a concentration ratio of 1.5:1. Deglycosylated p75^{NTR} was produced by using tunicamycin to inhibit glycosylation.

Crystallography. Crystals of the NT-3–p75^{NTR} complex were grown by hanging-drop vapour diffusion at 17 °C. X-ray diffraction data were collected at EMBL beamline BW7A (Hamburg) to a resolution of 2.6 Å, and were integrated and scaled with DENZO and SCALEPACK¹⁹. The structure was determined by molecular replacement with two models of human NT-3 (PDB code 1NT3) and the partial C-terminal half of p75^{NTR} from the NGF–p75^{NTR} 2:1 asymmetric complex (PDB code: 1SG1) using the program Molrep²⁰ from the CCP4 program suite²¹. The remaining structure was traced by Coot²² independently into the electron density map, which was calculated with partial phases. The structure was refined with Refmac²³ to an $R_{\text{work}}/R_{\text{free}}$ of 22.4%/28.8%. All images were prepared with Pymol²⁴ except Fig. 2d, which was generated with ESPript²⁵. Crystallographic data are provided in Supplementary Table 2.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 10 April; accepted 16 May 2008.

Published online 2 July 2008.

- Schweigreiter, R. The dual nature of neurotrophins. *BioEssays* **28**, 583–594 (2006).
- Bothwell, M. Functional interactions of neurotrophins and neurotrophin receptors. *Annu. Rev. Neurosci.* **18**, 223–253 (1995).
- He, X. L. & Garcia, K. C. Structure of nerve growth factor complexed with the shared neurotrophin receptor p75. *Science* **304**, 870–875 (2004).
- Nykjaer, A., Willnow, T. E. & Petersen, C. M. p75^{NTR}—live or let die. *Curr. Opin. Neurobiol.* **15**, 49–57 (2005).
- Aurikko, J. P. et al. Characterization of symmetric complexes of nerve growth factor and the ectodomain of the pan-neurotrophin receptor, p75^{NTR}. *J. Biol. Chem.* **280**, 33453–33460 (2005).
- Banner, D. W. et al. Crystal structure of the soluble human 55 kd TNF receptor–human TNF β complex: implications for TNF receptor activation. *Cell* **73**, 431–445 (1993).
- Mongkolsapaya, J. et al. Structure of the TRAIL–DR5 complex reveals mechanisms conferring specificity in apoptotic initiation. *Nature Struct. Biol.* **6**, 1048–1053 (1999).
- Butte, M. J., Hwang, P. K., Mobley, W. C. & Fletterick, R. J. Crystal structure of neurotrophin-3 homodimer shows distinct regions are used to bind its receptors. *Biochemistry* **37**, 16846–16852 (1998).
- McInnes, C. & Sykes, B. D. Growth factor receptors: structure, mechanism, and drug discovery. *Biopolymers* **43**, 339–366 (1997).
- Wiesmann, C., Ullsch, M. H., Bass, S. H. & de Vos, A. M. Crystal structure of nerve growth factor in complex with the ligand-binding domain of the TrkA receptor. *Nature* **401**, 184–188 (1999).
- Wehrman, T. et al. Structural and mechanistic insights into nerve growth factor interactions with the TrkA and p75 receptors. *Neuron* **53**, 25–38 (2007).
- Rodríguez-Tebar, A., Dechant, G., Gotz, R. & Barde, Y. A. Binding of neurotrophin-3 to its neuronal receptors and interactions with nerve growth factor and brain-derived neurotrophic factor. *EMBO J.* **11**, 917–922 (1992).
- Helenius, A. & Aebi, M. Intracellular functions of N-linked glycans. *Science* **291**, 2364–2369 (2001).
- Baribault, T. J. & Neet, K. E. Effects of tunicamycin on NGF binding and neurite outgrowth in PC12 cells. *J. Neurosci. Res.* **14**, 49–60 (1985).
- He, X., Chow, D., Martick, M. M. & Garcia, K. C. Allosteric activation of a spring-loaded natriuretic peptide receptor dimer by hormone. *Science* **293**, 1657–1662 (2001).
- Ullrich, A. & Schlessinger, J. Signal transduction by receptors with tyrosine kinase activity. *Cell* **61**, 203–212 (1990).
- Reichardt, L. F. Neurotrophin-regulated signalling pathways. *Phil. Trans. R. Soc. B* **361**, 1545–1564 (2006).
- Barker, P. A. High affinity not in the vicinity? *Neuron* **53**, 1–4 (2007).
- Otwinski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326 (1997).
- Vagin, A. & Teplyakov, A. MOLREP: an automated program for molecular replacement. *J. Appl. Cryst.* **30**, 1022–1025 (1997).
- Project, C. C. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **50**, 760–763 (1994).

22. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132 (2004).
23. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D Biol. Crystallogr.* **53**, 240–255 (1997).
24. Delano, W. L. The PyMOL Molecular Graphics System, DeLano Scientific, San Carlos, CA, USA (2002).
25. Gouet, P., Courcelle, E., Stuart, D. I. & Metoz, F. ESPript: analysis of multiple sequence alignments in PostScript. *Bioinformatics* **15**, 305–308 (1999).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank X. X. Yu for help with analytical ultracentrifugation assay, and Y. Y. Chen for the BIAcore assay. We thank Genentech for the gift of

recombinant human NT-3 and recombinant human NGF. This research was supported financially by the National Key Basic Research Program, the National Natural Science Foundation of China and the National High Technology Research and Development Program of China.

Author Contributions T.J. supervised the project. Y.G. expressed, purified and crystallized the complex and performed biochemical assays. P.C. determined the structure of the complex. H.J.Y. helped with its expression and purification. P.C., Y.G. and T.J. interpreted data and wrote the paper.

Author Information Atomic coordinates and structure factors have been deposited in the Protein Data Bank under accession code 3BUK. Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to T.J. (tjiang@ibp.ac.cn).

METHODS

Cloning and plasmid construction. A construct encoding the extracellular domain (residues 1–161) of p75^{NTR} was amplified from the rat p75^{NTR} cDNA by polymerase chain reaction and cloned into pFastbac1, which was modified by adding an N-terminal honeybee melittin signal peptide and a C-terminal hexahistidine tag. Recombinant baculovirus was produced and amplified with *Spodoptera frugiperda* (Sf9) cells in serum-containing *Trichoplusia ni* medium–Fred Hink (TNM-FH) insect medium at 27 °C.

Expression and purification of proteins. The Sf9 cells were cultivated in flasks in 2.5 l of serum-free HyQ SFX medium (HyClone) at 27 °C. When the cell density reached 5×10^6 cells per millilitre, the cells were centrifuged and resuspended in 2.5 l of fresh serum-free medium and were infected with recombinant virus at a multiplicity of infection of more than 5. The supernatant of the cultures was collected 72 h after infection. The expressed protein was purified by metal-affinity chromatography on chelatin resin and size-exclusion chromatography on a Superdex 200 10/300 GL column (GE, shown in Supplementary Fig. 1). Deglycosylated p75^{NTR} was produced with tunicamycin to inhibit glycosylation. Recombinant human neurotrophin-3 (hNT-3) and human nerve growth factor (hNGF) were expressed from *Escherichia coli* (a gift from Genentech). Murine NGF was obtained from the submaxillary glands of mice (AbD Serotec).

Crystallization, data collection and structure determination. For crystallization, the purified p75^{NTR} and NT-3 proteins were each concentrated and buffer-exchanged with 10 mM Tris-HCl (pH 7.5) buffer at a concentration of about 20 mg ml⁻¹. Crystals of the NT-3–p75^{NTR} complex were grown at 17 °C by hanging-drop vapour diffusion from a mixture of p75^{NTR} and NT-3 at a concentration ratio of 1.5:1. Initially, crystallization was performed under the conditions provided by the commercially available crystallization screening kits Screen I and II (Hampton Research); microcrystals were obtained in this way. After the reservoir solutions had been optimized, better crystals were obtained by mixing 1 µl of protein solution with 1 µl of reservoir solution and equilibrating against 1 ml of reservoir solution containing 0.9 M lithium sulphate, 0.05 M sodium citrate (pH 5.0) and 0.7 M ammonium sulphate. Block-like crystals appeared after incubation for 3–5 days and matured to their full sizes (typically, 0.3 mm × 0.2 mm × 0.2 mm) within 2 weeks.

Before data collection, the crystal was cryoprotected by being soaked briefly in paratone oil (Hampton Research) and flash-cooled in liquid nitrogen. X-ray diffraction data were collected to 2.6 Å by using EMBL beamline BW7A at 100 K and at a wavelength of 1.0030 Å. The data were integrated and scaled with

DENZO and SCALEPACK software. The crystal has space group R3 with unit-cell dimensions of $a = b = 125.8$ Å and $c = 133.1$ Å.

Crystal structure was determined by molecular replacement by using the program Molrep from the CCP4 program suite with two models of human NT-3 (PDB code 1NT3) and the partial C-terminal half of p75^{NTR} from the NGF–p75^{NTR} 2:1 asymmetric complex (PDB code 1SG1) as the initial search models. The phases obtained were improved by using RESOLVE²⁶, and the structure of the remaining part of p75^{NTR} was independently traced by Coot into clear difference Fourier maps and refined with Refmac; no non-crystallographic symmetry restraints were used in the refinement. The software packages Refmac and Coot were used to complete the model and refine it to a final $R_{\text{work}}/R_{\text{free}}$ value of 22.4%/28.8%. Statistics for data collection and refinement are provided in Supplementary Table 2. The quality of the final structure was evaluated with PROCHECK²⁷. A Ramachandran plot showed that most of the residues were in the favourable region (80.4%) and that no residues were in the generously allowed and disallowed regions.

Analytical ultracentrifugation. Sedimentation velocity experiments were performed on a Beckman XL-I analytical ultracentrifuge at 20 °C. Protein samples were diluted with PBS to 400 µl at a concentration of about 0.3 mg ml⁻¹. Data were collected at 60,000 r.p.m. (262,000g) every 3 min at a wavelength of 280 nm. Interference sedimentation coefficient distributions, $c(M)$, were calculated from the sedimentation velocity data by using SEDFIT²⁸.

BIAcore analysis. Real-time binding and kinetic analyses by surface plasmon resonance were performed on a BIAcore 3000 instrument (Pharmacia Biosensor AB). The eluent contained PBS and 0.005% Tween 20. Human NT-3 was immobilized on a CM5 chip by using an amine coupling kit, and the remaining coupling sites were blocked with 1 M ethanolamine (pH 8.5). Binding was evaluated over a range of p75^{NTR} (25–800 nM) or dg-p75^{NTR} (50–1,200 nM) concentrations at 25 °C. Kinetic parameters were further determined with BIAevaluation 4.1 software.

26. Terwilliger, T. C. Maximum-likelihood density modification. *Acta Crystallogr. D Biol. Crystallogr.* **56**, 965–972 (2000).
27. Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26**, 283–291 (1993).
28. Schuck, P. Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and Lamm equation modeling. *Biophys. J.* **78**, 61606–61619 (2000).